

US VS THEM

TRIBALISM, SELECTIVE EXPOSURE & HOW NON-PARTISAN
ONLINE NEWS MEDIA ENGAGEMENT CAN BE FACILITATED

BY
CHRISTOPHER DORAN

A THESIS SUBMITTED TO VICTORIA UNIVERSITY OF WELLINGTON IN FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF DESIGN INNOVATION

UNIVERSITY OF WELLINGTON

2019

ABSTRACT

Codeswitching is the action of switching between languages to better impart information to a recipient. This thesis introduces a set of codeswitching translator tools as a method of disrupting the potentially damaging structures of tribal politics through the manipulation of ideologically specific language norms. We first investigate how tribalism and group identity impact our ability to participate in political discourse. Using this insight from a host of different research disciplines, we design an iterative testing environment for a variety of 'codeswitching' translators in order to see the impact of translations ranging in complexity from simple word and syntax substitution through to machine learning back-translation. Though back-translation was not found to be an effective technique, simple substitution methods provided a foundation of effectiveness and proof of concept among test participants, especially those that identified as politically aligned.

CONTENTS

Abstract	2
Contents	3
Acknowledgements	4
Introduction	5
Chapter 1: Literature Review	9
1.1 Tribalism, Group Identity & Language	11
Papua New Guinea	11
Nerd Culture	11
Tribalism & Its Trade Offs	12
Modern Political Tribalism & The Partisan Divide	12
Language as a Weapon of Tribalism	14
Code Switching & Inter-Tribal Communication	15
1.2 Selective Exposure	17
Introduction	17
Relationship to Cognitive Dissonance	17
Buying a Car	18
Smokers	18
Selective Exposure & Politics	19
Results of Selective Exposure	21
1.3. Social Media Discourse	23
Introduction	23
The Internet as a Tool for Good Discourse	23
The Internet as a Tool for Bad Discourse	24
The Discourse Itself	25
1.4 All Together Now	27
Introduction	27
It's All Tribalism Man...	27
What this Means for Discourse	27
Can We Do Anything?	28
Chapter 2: Methods	33
2.1 Iterative Design Outputs & Testing	35
Introduction	35
The Survey Platform	35
Design Outputs & Testing	36
Survey 1	36
Survey 2	37
Survey 3	37
Additional Design & Output Testing	38
Sentiment Translator Survey	38
2.2 Hypothesis	39
Time & Accuracy	39
Tribal Alignment & Accuracy	39
Survey 1 & 2	39
Survey 3 - Machine Learning	40
Chapter 3: Results	43
3.1 Results	45
Introduction	45
Survey 1 Results	45
Survey 2 Results	46
Survey 3 Results	48
Additional Sentiment Translation Tool Survey Results	49
Results Overview	50
Chapter 4: Discussion	53
4.1 Discussion	55
Introduction	55
Tribalism & Language Norms Presenting in the Results	55
Reading Time & Accuracy	57
Selective Exposure, Cognitive Dissonance & Skim Reading	58
Applications & Rationale	59
Concerns & Moving Forward	60
Chapter 5: Conclusion	63
5.1. Conclusion	64
Works Cited	66

ACKNOWLEDGEMENTS

None, it's all me baby.

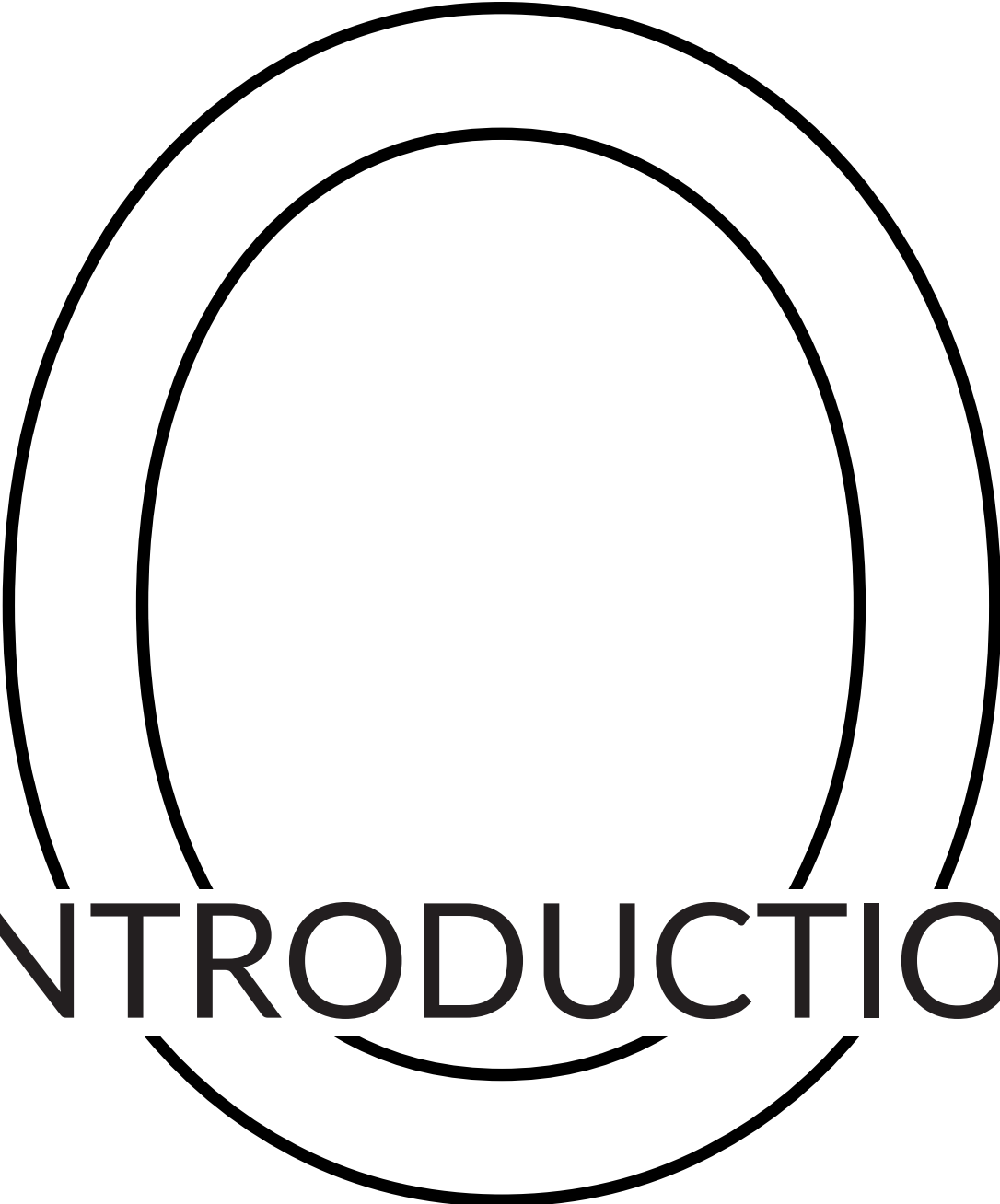
No, the fact that this thesis says "by Christopher Doran" on the title page is for the most part, complete claptrap.

This work could never have been possible without the unending patience and support offered by Tom White, who stuck with me with the determination of a curly haired, hawaiian shirt clad badger, while I repeatedly had him wait for me to get yet another superfluous portion of my body surgically removed. His wisdom and guidance, unlike Wellington's metlink website, never steered me wrong.

Another huge thank you to the amazing people from the MDI course. Stefan, Stefan, Phoebe, Bryan, Tess, Sophie and really everyone from that group. Thank you all for the spiral of motivational depression caused when you all graduated while I was writing my conclusion.

Another big thank you to my wonderful family, whose persistent ritualistic shaming and mental abuse helped me become who I am today. Just kidding, you guys were an endless supply of love, support and light in dark places, when all other lights seemed to go out.

I would also like to thank Laura, but I don't think I will.



INTRODUCTION

MODERN DISCOURSE & ITS PROBLEMS

Language is a key method for people to express their identity. Language can separate huge groups of human beings into countries and cultures, while also creating differences between regions within a country through dialects and accents. These social structures of effectively tribal groups that work to bind groups of human beings together under a social formation have been an important aspect of our species' survival throughout our evolution (Plater, 1990), (James, 2006). This security and social formation comes at a cost though, as the basis of many tribal structures is the creation of a group identity that people conform to and act out, in order to show their group membership. Intrinsic to the cementing of this identity is identifying and exacerbating differences between other tribes (Milroy, 1982). These structures remain throughout our cultures and societies today and it remains a natural element of human culture to build these tribal structures. Unfortunately, this can lead to difficulties in communication and discourse, as members of tribes can often value their membership to their tribe above the benefits of open and polite discourse.

The last 3 years have been the stage for some of the largest and most contentious political conflicts within recent memory. Between Donald Trump's historic election and Britain's referendum to leave the EU, modern politics has become a highly charged and divided environment. At the heart of the discussion of these topics is a growing divisiveness within political discourse, and the effects and future implications of modern digital communication mediums. Technological advances have consistently directly affected our methods of communication. From the invention of papyrus, to the printing press through to the creation of the internet, our forms of communication have become more inclusive and extensive than ever before. However, with this increase in capabilities comes the potential for misuse and abuse. Trolling, online bullying, fake news, mob mentalities and character assassinations have become frequent expressions of this globalistic freedom of communication. Among the problematic representations of this technology are the issues of the partisan divide, political bias and selective exposure (Zillmann & Bryant, 2013), (Hawkin, Yudkin, Juan Torres & Lynch, 2018). With hundreds of news outlets to choose from, individuals have the ability to pick and choose the news they wish to consume based on its impact on their pre-existing worldview or ideology. This is linked to the fact that within the last few years, US political party politics have become significantly more tribalistically motivated. Terminology such as "filter bubbles" and "echochambers" have been coined in order to describe the ways that social media users organise themselves tribalistically, and individuals are more likely to vote on a partisan basis than ever before, with split-ticket voting at the lowest rate observed since the early 70s (Resnick, et al., 2013), (Brownstein, 2016). Nestled within the rise in partisan division, increase in bias and the use of selective exposure is a question of whether our access to modern digital communication technologies are amplifying

these tribalistic instincts, and whether this can be avoided.

This thesis presents a set of design outputs and tests, culminating in proof of concept for a potential method of disrupting the way modern, politically active individuals consume online media and partake in online discourse. The ultimate, final output for this thesis being the proof of concept produced, built upon the set of iterative design tests, rather than several smaller outputs. Based on broad research, from political science and linguistics through to psychology and sociology, three iterative language tools were designed with the intention of utilising the power that language holds over tribal groups, by translating the language dialects of a sentence between ideologies. In this case, sentences were translated between U.S. Conservative and Liberal language norms. Each of these outputs were then tested through several surveys to discern their effectiveness at re-presenting politically charged information. The results show a foundation of effectiveness that can be built on with increasing complexity to use individual's tribalistic information processing methods against them, in order to expose these individuals to a wider range of political media. The results also point to the power that tribalism and reading style has over how individuals selectively expose themselves online, showing that the skim reading associated with many online mediums exacerbates the potential for individuals to screen information. Ultimately, the tools tested showed the potential to disguise information of one ideology in the tribalistic language norms of the other, in order to help disrupt the damaging structures of political tribalism.

1

CHAPTER ONE

LITERATURE REVIEW

1.1 - TRIBALISM, GROUP IDENTITY & LANGUAGE

Introduction

The division and conflict that prevails throughout our cultural landscape is no new condition, instead it is simply the latest development in a long and storied history of unassailable differences, tracing all the way back to when conflict drove us to hide in different caves simply to keep our foes out of our field of vision. In today's retelling of the tale, our caves are ideologies, insulated by endless access to tailor made information and media. Our tribe is no longer centered around one camp fire, but spread across the world, connected, not by adjacency, but by telecommunication, social media and the results of continued globalisation. Despite this fragmentation of the traditional formation of a tribe, the concept of loyalty to the ideals and beliefs of one's tribe above all else remains. At the center of this structure is language.

Papua New Guinea

Language is used by groups to form an identity, and to solidify the uniqueness of that identity by using it to highlight differences with opposing or differing groups. In pre-colonial Papua New Guinea, language was so tied with the idea of the tribe that linguistic groups were small, with neighbouring villages, short distances apart speaking distinctly different languages and dialects. These differences were championed and celebrated by the tribes. Intermarriage and awareness of differing languages is believed to have led to "heightened consciousness and pride in linguistic differences", rather than of integration and assimilation (Milroy, 1982, p 207). A high societal value was placed on bilingual individuals within this culture, as they represented the only avenue for communication and cohabitation between tribes. These individuals held a high standing in their societies as the key to conflict and conflict resolution. In this context social divisions and language divisions are intimately linked and related, with neither being the clear cause of the other, but instead a complex and deeply intertwined relationship. Division is created and perpetuated by language difference. This perpetuation leads to an ever increasing uniqueness among the tribes linguistic traits, thus creating a cyclical effect of difference creating more difference. Yet despite the collective quest for "pride in linguistic difference" the understanding of the importance and value of mechanisms to bypass these divides remains.

Nerd Culture

Fast forward hundreds of years and travel thousands of miles. At first glance the world could not seem more alien compared to the hilly island home of the Papua New Guineans. However despite distance and time, some structures and hierarchies repeat themselves in oddly recognisable ways. Within the typical US suburban school one will find a new set of individuals, each one aligning his or her identity to a group. According to Mary Bucholitz, you will find Jocks, Burnouts

and Nerds, each group employing a set of either positive identity practices or negative identity practises. Where the former is a set of actions to bond one further to his/her preferred group and the latter is a set of actions to separate oneself from competing groups. Bucholitz argues that the language norms created by, for example, the Nerds, are created to reject the qualities of the Jocks and the Burnouts and to use this difference to create a new group identity. This could be rephrased, possibly to the familiarity of the reader of this paper, as "a heightened consciousness and pride in linguistic differences." Hundreds of years and thousands of miles away from the pre colonial people of Papua New Guinea and still employing language dialects and modes of speech to create and maintain the structure of tribal groups. The similarities do not end there. Sociologist David Kinney argues that in order for "Nerds to succeed socially" they must undergo a process which includes "broadening one's friendship network, participating in extracurricular activities, and heterosexual dating," essentially displaying the qualities of a Jock. In other words, a greater value is placed on the members of one tribe that are "bilingual," in terms of linguistic and social traits, mirroring the status effects of bilinguality in Papua New Guinea. In many of these instances, tribalism and group identity practises have produced societal benefits. Tribalism acts as an "instinctive affiliation between people" that "has evolved to have remarkable staying power and utility" (Kysar and Salzman, 2003, p 1105).

Tribalism & Its Trade Offs

In many cases, a strong social unity has been an essential adhesive for the fabric of a society to remain intact. Tribalism can be a "precious cultural commodity", it helps "people survive in settings of extraordinary stress, deprivation, and complex antagonisms." It bonds the tribes of Papua New Guinea with pride in language and it insulates the tense social structures of the suburban high school. Human beings have evolved to make use of these 'social formations' as a defence mechanism and it remains a "strategic survival mechanism in much of the Third World" (Plater, 1990, p 5), (James, 2006). However, these formations and the security they provide come with a trade off. The tribes of Papua New Guinea lose the ability to communicate with other nearby tribes and high schoolers are subjected to ritualistic torment based on their group identities. The safety, cohesiveness and social adhesion of the tribal unit is created by drawing further attention to the differences that exist between other tribes. The result of this emphasis placed on creating, maintaining and advertising difference among groups leads to the values and interests of 'rival' groups being "excluded from recognition in the process of governance" (Plater, 1990, p 5).

Modern Political Tribalism & The Partisan Divide

Whether it has been the case for decades or centuries, or represents a new threat to modern democracy, the effects of tribalism can be seen in the apparent

partisan divide that prevails throughout modern political culture.

In modern politics, the narrative of conflict is dominated by two main groups. Conservative, and Liberal. These groups are “tribal in their behavior, emphasizing their group identity, ideology and the threat of the “other”” (Hawkins, Yudkin, Juan-Torres and Dixon, 2018, p 74). It is hotly debated as to whether the current US political landscape is more, or less divided than previous decades. Despite this debate, it is important to note that “split ticket voting is at its lowest on record” and that according to a recent study by the More In Common Initiative, within the highly politically active demographic, 50% of individuals do not believe that they need to be willing to listen to opposing viewpoints, but instead should “stick to their [group’s] beliefs and fight.” This implies that people are shifting away from an ideological middle, and are filtering closer to the extremes of partisan ideologies. It also shows that as people shift away from holding both conservative and liberal views to only holding one or the other and that they are growing less likely to listen or discuss opposing views without being combative. This is true in so many modern contentious topics, that in their article regarding political tribalism within the environmentalism and global warming debate, Kysar and Salzman discuss the possible importance of the “creation of a “radical middle” to break this tribal deadlock.” Their conclusion being that the framing of the debate as “two opposing philosophies and a vacuum in the middle” will not lead to action, but instead an increase in hostile discourse (Kysar and Salzman, 2003, p 1103). While political ideology has always been a natural division point for the formation of tribes or social formations, it is a recent trend that traditional electoral strategies “aimed at persuading swing voters” have been deemphasised and replaced with “an emphasis on maximizing turnout of the base” (Fiorina and Abrams, 2008, p 565). This is both caused by, and is causing the lack of swing voters in general. This change in tactics for political parties to increase support is an extension of the tribalistic tactic to “systematically [include] all members of [their] community and [exclude] all others as “outsiders” (Plater, 1990, p 2). However, it would be ironically biased of this thesis to not also discuss the counter theory that polarisation is an over diagnosed condition. This mode of framing the political landscape argues that the use of the term “polarisation” to define the idea of a difference of opinion is too broad. For example if we look at figure 1 and 2 together, it would be a quick and easy judgment to say that fig 2 is an example of a polarised community, with larger distributions on the left and right extremes, whereas within fig 1, the distributions are more centrally located therefore less polarised. However, it could also be argued that as “half of the respondents fall left of center and half right of center” in both examples, the term polarisation would apply. With this in mind, the studies that use the conclusion that polarisation is highly present within society as an argument that the political climate is fractured and highly divided could be called into question. That leads “skeptics of the polarization narrative [to] say that “fragmented” or “heterogeneous” [are] a more accurate descriptor[s]” (Fiorina and



[Fig 1]



[Fig 2]

Abrams, 2008, p 565).

This thesis intends to keep these semantic critiques in mind when examining the conclusions of polarisation based research, as Figure 2, not Figure 1, would be the definition of polarisation that this research concerns itself with. That being said, this critique of the broad definition of 'polarisation' and the implications of inaccuracy that definition could create, does not debate the argument of tribalistic tendencies, and the associated language trends within these groups, which is the intended focus of this exploration.

It would be plausible that if we were not partially inhibited by the connotations of 'civilisation' and instead were to look at society through the lens of someone viewing an ancient society, that an "anthropologist in Washington D.C. [could] discern the existence of [many cohesive tribes]" (Plater, 1990, p 6). Urban tribes and Farmland Tribes, 'Native' Tribes and 'Immigrant' Tribes, Black Tribes and White Tribes, Environmental Tribes and Deregulation Tribes, Pro-Life Tribes and Pro-Choice Tribes. Increasingly, membership to each of these tribes can be accurately predicted by one's alignment to one of two umbrella tribes/factions, Liberal or Conservative, Democrat or Republican. Each tribe, just like those of pre colonial Papua New Guinea, values their unique languages and practises as a source of pride and difference to the extent of alienating "outsiders." These allegences have the effect of emphasising narrow beliefs and privileges to the "net detriment of the nation as a whole." In the words of J.B. Plater, "this preclusive tribalistic inclination seems to increase the more that there are strains, adversities and limitations upon the resource system, in precisely those situations where increased integration and cooperation would appear to be most necessary" (Plater, 1990, p 6).

Language as a Weapon of Tribalism

These examples demonstrate that the concept of social formations, tribalism and group identity is an inescapable one. Where hundreds of years apart and in completely different community contexts, the same tropes of a quest for selective unity through separation appear. These same rules apply to hundreds if not thousands of other settings, from race, through to politics and all the way through to taste in media. It is also clear that central to the formation of these groups are the performances of positive and negative identity practices, including the use of language and dialect. When groups are divided over a common goal, it is natural for conflict to arise, including the creation of language specific division, such as derogatory terms and a group specific dialect (Sherif, 1958). Individuals create their verbal vernacular to help solidify their place within the group or social formation that they identify with. Milroy and Margrain mention a score of other researchers' ideas on how an increased adherence to a group's language vernacular norms is correlated with an increased integration in that groups hierarchical structure (Milroy and Margrain, 1980). This idea provides the basis for the concept that

reinforcement of currently accepted language traits, rather than an openness to the traits of other groups facilitates greater social standing within the individual's tribe. These norms serve to propagate difference in intergroup discourse and attempt to establish power distinctions (Milroy, 1982). In many instances, part of the dialect a group creates and disseminates throughout its members is either openly or subtly meant to force division and superiority through derogatory terms and phrases. Consider the terms and phrases used to invalidate or lessen the value of another tribes ideology Examples of this can be found in all walks of life, including racial and sexist slurs and exclusionary legislation. In terms of the modern political tribes, we have liberal phrases such as "cultural appropriation," the concept of "checking privilege" and the proclivity to reduce opposing viewpoints to racist, sexist or the result of being a member of a dominant identity group, for example "white privilege." Conservative approaches include the idea of "political correctness going too far," implications of ideas centered around protection of minorities being weak or "snowflakey," and phrases such as "feminazi" or "social justice warriors." Elements of liberal and conservative group identity are framed through these language norms, and membership to each tribe can be assumed or correlated to an individual's proclivity to use them. Given that social standing in a tribe is related to, and benefits from the propagation of these language norms and that these language norms often include phrases used to invalidate "outsider's" ideals, we can argue that the modern member of the liberal or conservative tribe is rewarded by using conflict based rhetoric to reinforce their self identity.

Code Switching & Inter-Tribal Communication

Similar to the societal value of bilinguality in Papua New Guinea, and the benefits of crossing through group identity in high school cliques it is reasonable to assume that there should be some social utility in having command over both ideological dialects in today's political groups. One possible avenue to explore in terms of exposing audiences to opposing ideologies comes in the form of linguistic research around persuasion, rhetoric and empathy. "Language changes can trigger social changes" and particular usages of words and phrases with rhetoric or ideological meaning attached can "trigger off certain behavioural patterns" (Wodak, 1989, p 95). In terms of the tribal communication this thesis is discussing, this could refer to the different dialects that both political tribes communicate using. These dialects create "behavioural patterns" in both parties. It is possible that shrouding a liberal message in the language norms of the conservative tribe could aid the reception of that idea for a conservative tribesman by playing upon these behavioural triggers created by the language pressures created by this tribalism. For example, a modern Conservative tribesman may see a message including a key word or phrase that they associate with their own tribe, thus become more likely to take the time to parse that message, given that their first impression is that it conforms to their tribalistic norms. This also ties in with Prof. Rob Willer's insights

into Moral Reframing, wherein ideals that are generally tied to one ideology are easier to digest by the opposing ideology when it is framed within a context that is familiar (Shashkevich, 2017). This idea of merging an ideology's message with an opposing ideology's language is not without precedent. Codeswitching is a linguistic term for bilingual speakers who switch between languages on the fly. For these people, switching between language becomes related to contextual changes (Woolard, 2007). Codeswitchers alter between languages with a "consciousness of typical associations of the language" they are using. For example, if speaking with an individual with the same language ability, a codeswitcher may swap between French or German based on which language will better impart the best way to cook a crepe. French may be best suited for a discussion on the tastes and smells that are desired, while German may be better used for the specifics of what systems work to make the best dessert.

In effect, they use the language that will best display their point to their audience; (Woolard, 2007) this is exactly what could be done with political discourse. By empathetically organising political arguments into the particular dialects common with the members of the opposing ideology, it is possible that political messages could breach the partisan blockade currently affecting the political landscape.

1.2 - SELECTIVE EXPOSURE

Introduction

Taste and identity are heavily linked. Through teenage years, and in many ways beyond, so much of our social status and place in society is decided by what media, clothes or sports we do or don't like (Kysar and Salzman, 2003). When our relationships and self identity begin to be influenced by the media, fashion or trends we like, an incentive to moderate our exposure to outside elements is created. This is partly due to how an expression of taste in any area of culture has the effect of being a form of tribalistic communication. The effect of this conformity being to strengthen one's standing in their chosen group. This implies that for many people, a portion of what media or trends they seek out is a reflection of their own tribalistic ideologies perpetuating themselves.

When narrowing this concept of taste and selective conformity to the idea of modern news and political media, it becomes clear that there is one phrase that underpins them all, one phrase that can identify, link and bind the discussion of modern political media consumption together, Selective Exposure. This being the idea that an individual chooses what media to expose themselves to based on what best fits their ideology and what will be most agreeable to them or their tribe. This action provides the individual a level of intellectual shielding from concepts that could threaten a sense of self identity. However the cognitive safety net provided also potentially comes with a level of insulation that may be the grounds for an increase in a polarised modern news media landscape.

Relationship to Cognitive Dissonance

While it is possible to frame selective exposure within the natural activities of a tribal group structure, instead in most cases this phenomenon is categorised as an extension of Leon Festinger's theory of cognitive dissonance. The basis of which is that when an individual holds several 'cognitions' that are opposed or inconsistent, 'dissonance' is created. The magnitude of this dissonance is heavily dependant on how important the cognitions are to the individual. This dissonance is an uncomfortable and undesirable status effect to hold, so naturally the individual will act to remove this discomfort, often by "moderat[ing] amounts of dissonance" (Zillmann & Bryant, 2013, p 12). Zillmann and Bryant's basic example of this concept is to imagine an individual with the cognition that they are a "good and truthful person" (Zillmann & Bryant, 2013, p 11). Then to imagine that person telling a lie. After telling the lie, the individual holds two competing cognitions, that they are a good individual and that they know they have just told a lie. This situation will create dissonance, and in order to reduce it, the individual may attempt to rationalise their lie in order to frame it within their preconception of themselves as a good and honest person. "I did it for their own good," they may say. "It was a small, white lie" they may implore themselves. "They deserved it," they might conclude.

Ultimately, Festinger proposed that people attempt to reduce the dissonance

created by competing ideologies or concepts by a range of coping strategies to protect their current and safe world view. When it comes to politics, there is likely no clearer set of conceptual ideals in greater competition than liberalism and conservatism. Leading to a huge quantity of individuals and groups seeking resolution and release from the resulting dissonance.

Buying a Car

Setting politics aside for a moment, like at a respectful family gathering, first it's important to frame the effects of cognitive dissonance and the use of selective exposure as a coping mechanism in a more neutral setting. Purchasing any expensive item is cause for reflection. Even more so when choosing between two nearly equally attractive alternatives. The choice between two, let's say, cars, of seemingly equal or comparable quality, creates a competing set of cognitions that can produce a level of anxiety. Once a decision is made, a second level of dissonance is created; was the Ferrari the correct decision after all? Or would the Tesla been a better pick? This dissonance is the birth place of selective exposure. Festinger claims that to combat this dissonance, the individual will begin to selectively expose themselves to future literature and media relating to these cars. Choosing to favour information that supports the decision, thus reducing and combating the anxiety and uncomfortable feelings relating to all the possible futures or missed opportunities created by choosing that damned Ferrari. In a 1957 study by Ehrlich, Guttman, Schonbach and Mills, the level of post purchase selective exposure was measured in new car owners. The results supported Festinger's claims, showing that owners would be more likely to read ads relating to the car they had chosen than ads of cars they considered buying, but opted against (Zillmann & Bryant, 2013).

Smokers

Leaving the car lot and stepping into the shoes of a smoker, we can see a new, yet familiar example of the dissonance to selective exposure feedback loop. The dissonance is created by the competing ideas of the desire to continue smoking vs. the knowledge of health related effects. Smokers exist in an often hostile media landscape, as anti-smoking imagery is hard to escape. However, where any degree of choice exists it would be reasonable to expect an individual to act to reduce the presence of their dissonance. A study by Brock and Balloun in 1967 tested this by exposing smokers to two sets of tape recordings. One set supporting smoking, the other opposed to it. While listening to the tapes, static would regularly obscure what the individual was listening to, forcing them to repeatedly press an antistatic button to clearly hear whatever message had been playing. In this instance, the smokers used selective exposure to combat the dissonance created by the competing ideas. They regularly pressed the antistatic button for supportive messages, yet tended not to use the button

on the anti-smoking messaging (Cotton, 2013, p 23), (Brock and Balloun, 1967). Interestingly, in a separate study out of the University of California, discussing the prevalence of smokers in adolescents found that the use of selective exposure also predicted a higher rate of smoking. This indicates a sort of feedback loop of ideological information exposure for individuals, wherein selective exposure both results from strong ideological positions and can lead to them (Sussman et al., 1988). In both of these examples, from buying a car to rationalising smoking, individuals can be seen to use selective exposure as a defense mechanism of sorts. Using it to combating uncomfortable feelings towards an idea that they hold dear, by insulating themselves from contradictory information.

Selective Exposure & Politics

For a short term solution this is effective, however in both situations, the selective exposure demonstrated seems to tend towards the limitation of self discovery or growth, instead of a willingness to broaden one's horizons. Decisions like these exist throughout all levels of society, from as simple as which game console to purchase through to who to vote for. Similar to the results of buying a car or smoking, people tend towards allowing these decisions to influence how they engage in discourse. Xbox vs. Playstation, Ferrari vs. Lamborghini, Smoking vs. Non-Smoking; some significantly more benevolent and harmless than others. In the realm of politics, particularly US politics, this takes the form of Liberal vs. Conservative. In this setting there are higher stakes involved with insulated and defensive thinking strategies as compared to the effects of buying a car. The same strategies that help a recent car buyer feel happy with their decision can lead to a fragmenting of traditional media institutions with regards to politics (Garrett, Carnahan & Lynch, 2011). In this way, tribalism and selective exposure are bonded in the mutual goal of facilitating one view point in opposition to others. Individuals are rewarded for insulated beliefs from both within and without, through the minimisation of personal dissonance and through group stigma (Hawkin, Yudkin, Juan Torres & Lynch, 2018).

The modern political media landscape provides a particularly suitable environment for the use of selective exposure to limit the effect of dissonance. The presence of the internet presents a new conundrum as it "makes passive exposure more difficult, and selective information seeking easier," while "algorithms [give] priority to certain types of content on social networks tend to foster expressions of moral outrage" (Valentino, Banks, Hutchings & Davis, 2009, p 1) (Hawkin, Yudkin, Juan Torres & Lynch, 2018, p 133). This ease and algorithmic aid combines with the lower production costs of news media and the profitability of creating a consistent, partisan viewership to create a news environment that has shifted towards rewarding bias, rather than combating it (Baum & Groeling, 2008), (Bernhardt, Krasa, & Polborn, 2008). This concept

ties in neatly with the tribalistic reward for rewarding difference and separation from other tribes as a statement of one's tribal identity. For example, in a recent study on the 'Hidden Tribes' of US politics, researchers found that 72% of "Traditional Liberals" found the news network MSNBC to be honest, whereas 74% of "Devoted Conservatives" believe it to be "very dishonest." The same partisan relationship can be observed when referring to Fox News, as a vast majority of conservative individuals consider it to be honest, while liberals consider it dishonest. The researches then extended their analysis to find that each political ideology group thought it was incredibly likely that the news media of the opposing ideology would "report false stories if it benefits them" (Hawkin, Yudkin, Juan Torres & Lynch, 2018, p 133). This "defense motivation" to opposing ideas is a key tenant of both selective exposure and tribalism (Hart et al., 2009). In these cases, the discrediting of opposing news sources is a cure to dissonance, a rationalisation of selective exposure and an action of tribalistic identity practises all at once.

However, it would be a microcosm of what this thesis attempts to resolve to not mention the existence of a counter argument to this concept. Many researches have conducted studies that provide conversely concluding trains of thought. Conclusions vary between the idea there is little to no evidence of an intrinsic defense response to dissonance, that there is an element of selective exposure, but it is not the only predictor of information reception and that selective exposure is not ideologically charged, but instead motivated by an individual's anxiety. For example, building on Sears and Freedman's original 1967 critique of Festinger's conclusions a 2009 study found that when confronted with a defense motivation, or a concept that produces a level of anxiety or tension, individuals are more likely to seek out useful information, rather than ideological information. This concept diverges from Festinger's claims as it postulates that an individual will prioritise usefulness over their confirmation bias (Valentino, Banks, Hutchings & Davis, 2009). A similar, but slightly divergent conclusion was also reached in 2009 by a separate study that determined two motivators for information exposure, "defense" and "accuracy." Where individuals with defense motivation's were likely to selectively expose themselves to congenial information, whereas accuracy motivated individuals were not (Hart et al., 2009). Sears' original critique of Festinger in 1967 also included research on smokers that directly contradicted the research finding smokers to seek agreeable information. Instead it found that smokers were exposing themselves to unsupportive information, but then acting out their bias by "[subjecting] it to careful and mercilessly unsympathetic scrutiny" (Sears & Freedman, 1967, p 213). However, interestingly, in their article critiquing Festinger's theory, Sears and Freedman admit through a meta analysis of many other studies that selective exposure does seem to present in certain situations, just not generally and across the board. Instead they determine that there are

certain situations that seem to promote the use of selective exposure. Included in those is when an individual has just made an important decision or committed to an action and when an individual has been exposed to competing information against their will.

With this research I intend to operate with all of these conclusions in mind (Sears & Freedman, 1967). However, this thesis does not intend to settle the debate of whether selective exposure is generally presented. Instead it intends to test the tribalistic tendencies around information exposure surrounding language and dialect usage with a secondary emphasis on areas of extreme partisan fragmentation wherein selective exposure has been acknowledged to a higher degree than less intensely dissonant examples. I also argue that selective exposure is not always an inward psychological trait, but is instead also a required outward expression of tribalistic identity, and a mostly unavoidable phenomenon due to external factors, including information distribution algorithms and news outlet bias'.

Results of Selective Exposure

"Exposure to a diverse marketplace of ideas is a central tenet of deliberative democracy" (Garrett, Carnahan & Lynch, 2011, p 1). This thesis argues that while the internal and external motivators for selective exposure have always been present, whether they have resulted in selective exposure or not, the modern media landscape plays a key role in facilitating a level of freedom in selective exposure that has not been possible before. The fragmented and ideological nature of modern news media leads to a level of ease when selectively exposing oneself to information. While selective exposure would not specifically create an uninformed populace, it would create a populous that is informed of different topics. News items that benefit a particular ideology are often resigned to being shown only on networks or outlets that aim to promote that ideology, which means that when attempting political discourse, individuals are effectively missing pieces of information that can bridge their partisan divide. "Sources bias the news by suppressing or de-emphasizing certain events that could be perceived as unfavourable by their respective audiences" and due to this, liberal audiences are more likely to be informed of the benefits of their own candidate, and the negatives of the opposing candidate. Conservative audiences being the opposite (Bernhardt et al., 2008). This leads to a lack of knowledge in both camps about their own party leader's flaws and the benefits of the opposing party. This structure of information delivery and reception leads to an inherent division of knowledge, a division that is drawn along political partisan lines. Due to this, "important information is lost through [this] bias and can lead to electoral mistakes" such as the wrong candidate being elected (Bernhardt et al., 2008, p 1). This attitude towards news media on behalf of both outlets and audiences "may have serious consequences on the democratic social system" as it causes a

“fragmentation” within groups regarding political tolerance (Kobayashi & Ikeda, 2009, p 1). This fragmentation of discourse may initially seem like something superfluous to real life political movements, but in fact this change in political language towards a more divisive and conflict centered communication is actually contributing to a “trend that has defined voting behaviour for most of the past decade and that has left the two major parties increasingly homogenized and partisan” (Chandler, 2013, p 1).

1.3 - SOCIAL MEDIA DISCOURSE

Introduction

Internet technology seems to be the closest our modern advances have come to embodying the bold visions of the future that science fiction films of yesteryear artfully proposed. Now, in 2018 we are 3 years beyond the world's deadline for a hoverboard, as set by *Back to the Future*, and with only 1 year left to create the dystopian vision of the original *Blade Runner*, we are running out of time to explore Mars, create androids, flying cars and an Asian cultural infusion with modern day Los Angeles. This might sound like the futures we were promised by movies were false visions, but that would be incorrect. Today we can sit in our *Total Recall* (1990) self-driving car, browsing our *Space Odyssey* (1969) tablets on Mark Twain's web-like, interconnected global communication 'teleelectroscope' (1910), while berated by *Minority Report's* (2002) targeted ads and surveillance. Information and communication have never been more easily imparted. Billions of people hold access to near limitless data and communication opportunities. The internet in itself brings together some of the ideas considered to be beyond our reach 10 or 20 years ago under one roof, along with countless new and hitherto undreamt of gadgets and life-changing dohickies. As with the division of the science fiction genre between hopeful, speculative pulp and dystopian critiques on society, there exists a similar division in the academic understanding of the effects of the internet on social political discourse. An old familiar debate between a hopeful understanding of human nature, vs a critical take on our dark and primal tribal urges. Does social media create open and accessible debate? Or does it instead facilitate closed mindedness, increased tribalisation and conflict oriented discourse? Considering current research, it seems that the truth lies closer to the latter, though with a healthy measure of the former.

The Internet as a Tool for Good Discourse

Historically speaking, freedom of information and communication has been a key tenant of seeking social equality and moving away from oppressive regimes (Brundidge, 2010). With the discussion of social media's effects on public discourse, it is important to understand that the tools offered by this technology do hold the power to challenge power dynamics and empower citizens. For example the incredible power social media demonstrated during the Arab Spring revolutions (Huang, 2011). *Time Magazine* also recognised this global trend of the growth of individual power in the interconnected age by breaking the trend of using a world leader or power figure as their person of the year first in 2006, then again in 2011. In 2006, the person of the year was "You," a mirror in the magazine reflecting the face of the reader. Similarly in 2011, "The Protestor" found its way to the front page. Among other implications, these examples constitute a growing celebration of the power of the individual that may not have been possible without the notion of social media (Lance Bennett, 2012). With this

understanding it becomes easy to empathise with the hypothesis of the research into social media discourse that the use of the media is key to creating a more inclusive public sphere of political debate and societal betterment. In 2010, the Netherlands' politicians were twice as likely to have a twitter account than their British counterparts. Researchers investigating the impact of the quick adoption of twitter among the Dutch politicians found that 3 usage patterns presented themselves within the candidates twitter platforms. Firstly, politicians could use their account merely as information delivery systems, with little to no interaction with their base. However the other two usage patterns, coined as the "innovator" and "mobilizer," were found to afford "citizens the prospect of representative closeness, mutuality, coherence and empathy." They were also tied to more 'grassroots' campaigns, civic engagement and mobilization (Graham et al, 2016, p 16). It is hard to argue with the utility of social engagement and the ability to shift the power of information dissemination away from state bodies and into the hands of an engaged populace (Gaines & Wagner, 2013).

The Internet as a Tool for Bad Discourse

What is clear is that when it is working to its greatest potential, social media holds a great power to elevate the status of those who have been underrepresented, and to facilitate the civic engagement of a greater 'public sphere.'

Despite the clear incentives for offering an ease of public engagement, and the societal benefits of a greater access to free speech, there are elements of the opportunities social media can offer that are not conducive to positive discourse. Selective exposure, tribalism, algorithms and a tendency towards extreme sentiment have all arguably led to a fractured political discourse framework (Conover et al., 2011). From the meteoric rise of the term 'fake news,' to recent revelations of the less than innocent effect of spam/bot accounts, through to the growth of tribalistic ideologic group structures. Research framed in this light becomes reminiscent of dystopian science fiction. Telling a tale of a wonder technology, spiraling from miraculous for democracy, to undermining it. Instead of a tool to elevate the underrepresented, and change the power dynamics of information dissemination, some research instead suggests that social media, and twitter in particular serves to allow us to more efficiently act on our urges to combat dissonance and identify tribally. Included in this new toolkit is the addition of content sorting algorithms, that can tend toward creating the effects of selective exposure without an individual even actively attempting to negate dissonance (Bakshy, Messing & Adamic, 2015). This means even without any intention to, subconscious or otherwise, most social media users encounter a level of enforced selective exposure at the hands of the online platforms themselves (Resnick et al., 2013). The extension of tribalism and selective exposure into the realm of social media is so pervasive that users' political alignment can be predicted to an accuracy of 95% by analysing the "segregated community structure" of political ideology information dissemination groups on twitter (Conover et al., 2011, p

1). These ideologically bonded social media communities have been referred to as 'filter bubbles' or 'echochambers' due to their nature of excluding competing ideas and the repeating/promotion of similar, tribalistically inclined thought patterns (Resnick, et al., 2013). The existence, and prevalence of these ideological, tribal groups, aided in part by the algorithms inherent in the structure of social media communication, throw a significant cog in the works of the 'public sphere' theory. Indicating instead that social media may be in part culpable to a further fragmentation of political discourse.

The Discourse Itself

Intertwined with the structure of online communities are the types of communication used. Social media offers an interesting environment in this regard, as conversations occur across distances and often with a level of anonymity that would have been alien before the arrival of the web. This combined with algorithms that prioritise levels of engagement and that users "tend to organize into insular, homogenous communities segregated along partisan lines" has created an interesting ecosystem to dissect the influences of selective exposure and tribalism (Conover et al, 2011). Twitter is unique in that it is a network where users both use selective exposure to moderate their information input while also creating content that continues the cycle of selective exposure and partisan division in others. Add to that the ability to cause conflict between ideologies and you create a self-sufficient cyclical 'echochamber' for the continued creation and strengthening of 'ideologically segregated communities' (Conover et al., 2011). The problem with this medium is that for information to truly be propagated throughout any given network it needs to capture its audience. Capturing an audience on twitter requires a tweet's sentiment to be extreme in one way or another. This means that the information that audiences are most often exposed to are extreme in nature. There are two types of user interactions on twitter, user to user mentions, where twitter users direct public messages at each other, and retweets, where people promote a person's tweet by showing it to their own network. These two types of interaction showcase two very different types of political discourse. The retweet network is highly politically 'polarised on a partisan basis' (Conover et al., 2011). For a person to retweet a tweet, it must first be ideologically aligned with the person retweeting it. Within this interaction, it is also beneficial for the popularity of the tweet for it to have an extreme sentiment attached to it, either positive or negative (Stieglitz & Dang-Xuan, 2012). Statistically, these are the tweets that are exposed to the most users. That means that twitter user's political exposure mostly relies on hearing more and more extreme versions of ideas that they already agree with. This interaction model changes in regards to user-to-user mentions on twitter. User-to-user mentions do not have the same politically homogeneous network that retweets do, instead mentions more often than not cross political ideological boundaries. While this may seem like it could be constructive conversation, the concept that extreme sentiment in tweets equates

to their popularity remains relevant between retweets and mentions. This means that the tweets that cross partisan division are likely to use divisive and extreme language to generate conflict, rather than constructive discussion. Interestingly, this conflict-oriented cross-ideological engagement actually serves to strengthen the tribalistic groups. Causing them to insulate themselves further (Conover et al., 2011).

1.4 - ALL TOGETHER NOW

Introduction

Selective exposure, tribalism, polarisation/fragmentation, sorting algorithms, filter bubbles and echochambers are all highly interconnected and inseparable cocktail ingredients coming together to create a radical shift in modern political discourse. Nestled within all of these concepts is language. How we use it to limit our exposure, to identify ourselves, to create difference and how our access to language is curated beyond our control. Language is a key method of expressing tribalistic tendencies, and in this case the natural tribalistic tendencies of modern audiences, coupled with their attempts to mitigate dissonance is to a large extent contributing to the prevalence of filter bubbles, echochambers and the conflict oriented discourse found on social media.

It's All Tribalism Man...

Each of us contains the ancient capacity for tribalism, and this capacity is exerted in strange and powerful ways throughout our seemingly modern and advanced societies. A partially unavoidable aspect of our nature that once protected us, now serves to pull and drive pockets of our national and global groups apart (Van Bavel & Pereira, 2018). This thesis argues that much of the social media activity we see as detrimental to good discourse originates with the concept of tribalism and group identity. Some aspects such as sorting algorithms present of social media platforms are tangentially linked to tribalism, as they likely exist to best capture audiences and a method of doing so is to pander to tribalistic tendencies within customer markets. However, concepts such as selective exposure and cognitive dissonance, while traditionally studied separately from concepts of tribalism, have been seen to be a more directly "reflect [the] tribalism inherent to politics" (Skoric, Zhu & Lin, 2018, p 4). Regardless of whether a user's selective exposure is tribalistically motivated or if it is solely motivated by the mitigation of dissonance, the effect is the same. It drives people's exposure to news and discourse towards a more insulated and protected approach. Ending in a position that greatly benefits the innate and primal tribalistic urges. Ultimately the tribalistic trait that underpins and ties together all of these concepts is that of our "defend and aggress" response. In simplest terms, our "evolutionary tendency" to push perceived aggressors away, while solidifying our network of allies. This mentality can be traced through all of the concepts I have so far raised regarding the causes of our fragmented political discourse (Hobfall, 2018, p 1).

What this Means for Discourse

Anecdotally, it could be said that the effects of tribalization have been visible throughout the 2017-2018 political landscape, however it has also been more formally discussed and tested that the modern media bias, and selective exposure leads to a gap in knowledge between voting groups (Bernhardt, Krasa

& Polborn, 2008). This gap in knowledge provides a difficult problem for polite bipartisan discourse to overcome. Individuals that identify as either party, for reasons relating to selective exposure and tribalism, are exposed to a different set of news items than their opponents. This means that when engaging in debate, these groups approach each other with large gaps in their knowledge regarding their opposition's position. This is unlikely to lead to positive debate, as both parties are often only knowledgeable about the negative aspects of each other's tribes and more likely to lend itself towards conflict oriented discourse, which as discussed previously, can actually result in a deepening of ideological, tribalistic divisions (Conover et al., 2011). Bernhardt, Krasa and Polborn argue that this type of media exposure and the resulting effects are likely to cause 'electoral mistakes.' What they mean by this is that because each party are only being exposed to the positive information about their own political tribe, and the negative information about the opposing political tribe, it is likely that their electoral decisions are based on misleading and incomplete information (Bernhardt, Krasa & Polborn, 2008). A more complete and accurate structure would be for both tribes to be exposed to each other's news to have a more complete understanding of the available information and possible discourse.

Can We Do Anything?

The conclusions regarding selective exposure and partisan fragmentation of social media users are difficult not to find reminiscent of the tribalistic tendencies of the Papua New Guineans and the dissonance combat techniques of smokers or new car owners. The same mechanics of insulation, difference and conflict are used to protect and retain an identity, and protect an ideal within ideological social formations. If one of the elements of remaining party to a tribe is to succumb to its language norms, it begs the question of if these natural language norms can be subverted to undermine these tribalistic structures.

If we can see each of these phenomena as a mode for which tribalism can exert its ideological and identity-forging force, we can also presume that the key tenet of language remains a constantly important signifier of tribalistic intent. With this in mind, it could be possible that the societal benefits for bilingualism may be available but underutilized in today's political media landscape. Codeswitching and bilingualism was lauded and essential throughout the examples of Papua New Guinea and the high school microcosm. It is possible that that same principle that values bilinguality, whether consciously in the case of Papua New Guinea and the cultural capital gained by being bilingual, or unconscious in the case of the high school, where social standing was improved by the adoption of traits associated with different groups, may also have a similar effect if applied to the tribal conflict of democrat vs republican. So, with that assertion in mind, and given that there seems to be little evidence of a growing contingency of modern political groups becoming or practicing bilinguality, could it be possible to create a programmatic approach that can make use of the tribalistic language dialects, selective exposure

and these echochambers to undermine the defensive structures of social media tribes? Would it be easier for a tribalistic, conservative social media user to read a liberal idea, if it was presented to them after the idea has been 'translated' to reflect conservative tribalistic language norms?

This thesis proposes that a possible avenue to undermine the current social media discourse structure is to use machine learning natural language techniques to translate ideologically charged information into the language norms associated with the the inverse or competing ideological group. In other words, taking a liberal leaning headline and codeswitching elements of that headline to be closer in line with what a conservative individual might consider part of their tribal dialect, may partially limit the dissonance created by reading the text, thus curbing that individual's need to use selective exposure as a coping mechanism. Techniques such as this have been used and studied in advertising environments, testing how bilingualism, dialects and slang can be used to increase the audiences impressions of a product or ideal. Similarly, historical examples of propaganda have seen their biggest successes when it relies on manipulating the core ideals, fears and norms of tribalistic groups.

Machine learning and natural language tools can be used to isolate key words, and language trends throughout large corpus' of data, then use those insights to convert text to more resemble the trends the algorithm identifies in the corpus of data. A tool showcasing one of the most successful versions of this new technology has already been tested on corpus' of liberal and conservative data, with successful translations occurring. The tool has been tested in a 2018 study to outperform some of the other most advanced forms of style transfer and machine learn translation.

Thus we propose to build set of iterative natural language processing tools, culminate in utilising this machine learning style transfer technique, along with developing relevant testing systems, with the purpose of testing the effectiveness of this programmatic bilingualism on political statements to various degrees of complexity. Taking advantage of tribalistic dialects in order to re-present information between tribes in a form that could avoid the information being dismissed on the basis of selective exposure, cognitive dissonance or tribalistic conformity.

[Right] Example translations of machine learning tool.

republican → democratic		
<i>i will continue praying for you and the decisions made by our government!</i>	<i>i am proud of you and your vote for us!</i>	<i>i will continue to fight for you and the rest of our democracy!</i>
<i>tom, i wish u would bring change.</i>	<i>i agree, senator warren and could be.</i>	<i>brian, i am proud to have you representing me.</i>
<i>all talk and no action-why dont you have some guts like breitbart</i>	<i>and then we will be praying for them and i am proud of this position and i am proud of</i>	<i>keep up and dont know, you have a lot of respect as breitbart</i>
democratic → republican		
<i>as a hoosier, i thank you, rep. visclosky.</i>	<i>a lot , i am proud of you <unk>.</i>	<i>as a hoosier, i'm praying for you sir.</i>
<i>thank you for standing up for justice and against bigotry-racism, homophobia, sexism, misogyny, religious and xenophobia.</i>	<i>do you for standing up for highly and in bigotry-racism, programming, cut, granddaughters, unprecedented and excludes.</i>	<i>thanks for standing up for the constitution and get rid of obamacare, homophobie, cut, and actuality.</i>
<i>thank you for all you are doing for us, attorney general harris!</i>	<i>thank you for standing up for us and i am proud of us!</i>	<i>thanks lawmaker for all you do for us, senator scott!</i>



2.1 – ITERATIVE DESIGN OUTPUTS & TESTING

Introduction

This thesis intends to discover what design tools can be used to undermine the pervasive effects of cognitive dissonance in online media and micromessages. To do this, a series of iterative versions of a translation tool have had their effectiveness tested through a set of multiple choice surveys. 3 surveys were crafted, each one composed of a selection of outputs from 3 variations of a translation/codeswitching language processor. Each survey will prompt a respondent to view 10 separate short statements, one at a time, and identify the statement as one of the following: Liberal, Conservative, Neutral or Unknown. Some of the statements that respondents view will be statements taken directly from a variety of sources on the web. Sources include news headlines, individual sentences from articles, through to comments and replies on social media platforms. Some of these statements will appear exactly as they did in their original online form (original statements/original messages) and some will have been translated using a natural language tool (translated statements/translated messages/new messages). The purpose of these translations is to attempt to change the perceived political leaning of the message to whatever ideology is opposed to the original message, while retaining as much of the message's original meaning as possible. For example, a translation applied to the original statement "Hillary Clinton creates great job opportunities," would intend to change that sentence from a potentially liberal leaning to instead seem conservative, while also attempting to maintain the underlying message that Hillary Clinton is good for the job market as much as possible. The types and variations of translations differs from survey to survey, as each iteration of the translation tools are intended to focus on a different type of linguistic change. These can range from simple keyword substitution, smaller syntax and grammar changes through to more extreme changes to the entire sentence structure.

The Survey Platform

The testing of these statements required a unique environment and set of tools. The first hurdle was that the survey needed to display only a selection of the available statements to a participant. Out of a larger pool of statements, only 10 were to be displayed, one at a time. This requirement was created due to the fact that the original statements that some translated statements are based on are also included as statements a participant could be exposed to during a survey. These original messages were included to act as a control variable. However, if a participant was to see the original statement prior to then also seeing the translated statement based on that same sentence, then it is likely that the results would be adversely effected or contaminated. The surveys also required each message to be displayed individually, with no other text on screen (other than the answers) and that the time taken to respond to each of the 10 statements was

recorded. While some survey platforms currently available offer elements of each of these requirements, the process of adapting this research's requirements to fit the predefined constraints of these other platforms led to sacrifices either in accuracy or efficiency. To combat this, I constructed a survey site specifically tailored to the needs of this research. Translated messages would never be displayed along with the original messages they were based on, statements were displayed one at a time, there was no superfluous text of any kind and the time taken to respond to each statement was recorded individually (Pictured Below). Designing and developing a specialized survey tool led to an increase in efficiency and allowed for an approach that benefitted from not needing to fit within the constraints of a pre-existing generalised tool.

Sen. Clinton actually promises more jobs

Liberal

Conservative

Neutral

Don't Know

Design Outputs & Testing

Survey 1

Within survey 1, translated messages from the first iteration of the language processing tool are put to the test. This iteration of the translator operates by isolation/selecting keywords or phrases that seem to lean towards a particular ideology and substituting these with words or phrases with the same overall meaning, but different ideological associations. For example, the noun "Hillary," could be substituted for "Lying Hillary." Similarly "President Trump" could be replaced with "Suspected Tax Evader Trump." While these changes have extremely different connotations, in most cases, these changes will leave the remainder of the message unaffected.

Liberal Version	Conservative version
Hillary Clinton, Hillary, Clinton, etc	Lying Hillary, War Criminal Clinton, etc
Obamacare, Socialised Healthcare, Affordable Healthcare Act	Communist Healthcare, Healthcare Socialism, Failed Obamacare, Trumpcare

[Example Translation Index]

Translation Process Example

- Step1 : Pick Statement
- "Hillary creates job opportunities"*
- Step 2 : Search Sentence for keyword or phrase matching available data
- "[Liberal Keyword] creates job opportunities"*
- Step 3 : Replace Liberal Keyword with one o f the Conservative Translation options
- "[Conservative Translation] creates job opportunities"*
- Step 4 : Output translated statement
- "Lying Hillary creates job opportunities"*

The incredibly basic and blunt linguistic changes used within this translation tool are intended to discover how effective the bare minimum of language substitution is with causing a change to how participants categorise a statement. The changes were limited to 1-2 translations per statement, in order to preserve as much of the original phrase as possible, to fully test the effects of the most minimal changes on a statement. Key elements recorded during this survey include: The participant's self identified political ideological standing, the participant's categorisation of each of the statements and the time taken to respond to each statement.

Survey 2

Like survey 1's attempt to discover the effectiveness of blunt, minimalist, relatively programmatically simple automatic substitution operations, survey 2 attempts to explore the effectiveness of more subtle changes to syntax and grammar. Where survey 1 operates by substituting whole words and phrases, survey 2 using a version of the translation tool that targets more subtle areas of sentence structure. This includes de-capitalisations, adding quotation marks, purposeful misspelling and removing full stops etc. For example "Hillary Clinton" could become "hilary clinton" and "President Trump" can become "'President" Trump." This could provide further insight into how small changes to sentences can impact participants perception of the ideological alignment of messages. Key elements recorded during this survey include: The participant's self identified political ideological standing, the participant's categorisation of each of the statements and the time taken to respond to each statement.

Survey 3

The third and final survey steps away from the smaller, isolated changes made by the translation tools used in the first two surveys and instead makes use of an algorithmic, machine learning approach developed by (source). In this iteration, the implementation of the translation tool is trained on one of two corpus' of text, Liberal or Conservative. The tool then takes an original statement and processes it using back translation to produce a sentence or phrase resembling the language norms demonstrated in the training text/corpus. The translations created by this process should represent a more holistic and in depth translation of the supplied text, and could result in a more nuanced resulting translated statement. This process will offer a comparison between one of the most advanced forms of machine learning back translation, and the more simplistic measures applied in previous translation tools and surveys. Key elements recorded during this survey include: The participant's self identified political ideological standing, the participant's categorisation of each of the statements and the time taken to respond to each statement.

Additional Design Output and Testing

For the purposes of further testing the effectiveness of substitution based modification of a sentence, a tangential design output was created. This additional output was made to test how simple word substitution could affect sentiment in a sentence without affecting the overall meaning of the sentence. Using Google's sentiment analysis API and a dictionary API, the translator analyses a sentence, discovers the keywords that are predominantly responsible for creating a negative or positive sentiment, and exchanges those words for a synonym that registers the opposite sentence sentiment. For example, the phrase "They are fat and ugly" could be analysed. The tool will determine that the sentence overall has a negative sentiment. Then the tool will discover that the two words most to blame for this negative sentiment are "fat" and "ugly." The translator will then search through synonyms of each of those words, looking for the words most positive sentiment, and place those words into the original sentence. In this case, the phrase becomes "They are plump and homely." Overall this sentence still describes a person as large and unfortunate looking, but the tone has become significantly less hostile. The purpose of this extra experiment is to further test how effective simple programmatic translation/substitution is when attempting to change people's impressions of a statement. It can also show the power that individual keywords have to colour a sentence's tone and connotations.

[synonym]ous

Enter a Twitter handle or a block of text that you would like to change.

You sneaky ugly stupid cat

Find tweet Change me

You sneaky ugly stupid cat

You sly homely dense cat

Sentiment Translator Survey

To test if this translation system can successfully change people's opinions of a statement, the survey system used for the political translators was adapted. In this variation responders were asked to select the tone/sentiment each statement showcased. Responders could select from: "Positive," "Negative," "Neutral" and "Unknown." Other than these changes, the other aspects of the survey remained identical to the system employed for previous surveys.

2.2 – HYPOTHESIS

Time & Accuracy

The design outputs with the hypothesis that testing the effectiveness of each of these iterative translators will ultimately reveal a linear relationship between time spent reading a particular statement, and the rate translated statements fail at disguising their original ideological leaning. For example, this thesis proposes that a statement that has been translated from an originally liberal sentence, to a faux conservative statement, will be more likely to be identified as liberal (bypassing the ideological translation) by participants that spend longer reading and studying the sentence. Similarly, participants that read and respond to sentences quickly will be more likely to be tricked by the translated version of one of these statements. Overall, it is likely that participants accuracy levels, being the rate at which they correctly identify the original meaning of a statement (or simply the only meaning in the case of an unchanged statement), will be increased with the amount of time spent on each survey question.

Tribal Alignment and Accuracy

Additionally, generally speaking it is likely that participant's tribalistic, political alignments will hinder their accuracy ratings when approaching translated phrases. A key result of tribalistic language norms is an individual's ability to quickly identify speech and written patterns of communication as either tribalistically aligned, or opposed. If this tenant of tribalism applies in these tests, participants will find it easy to correctly identify the political leaning of unchanged statements, but will struggle in correctly identifying any translated statements, as the more politically tribalistic the individual is, the more likely they are to be affected by the change to language norms. However, similar to the first prediction, an increase in time spent on the surveys is also likely to aid in mitigating the effects of participants own political leanings in their responses to statements.

Survey 1 & 2 - Small Changes, Keyword Substitution & Syntax

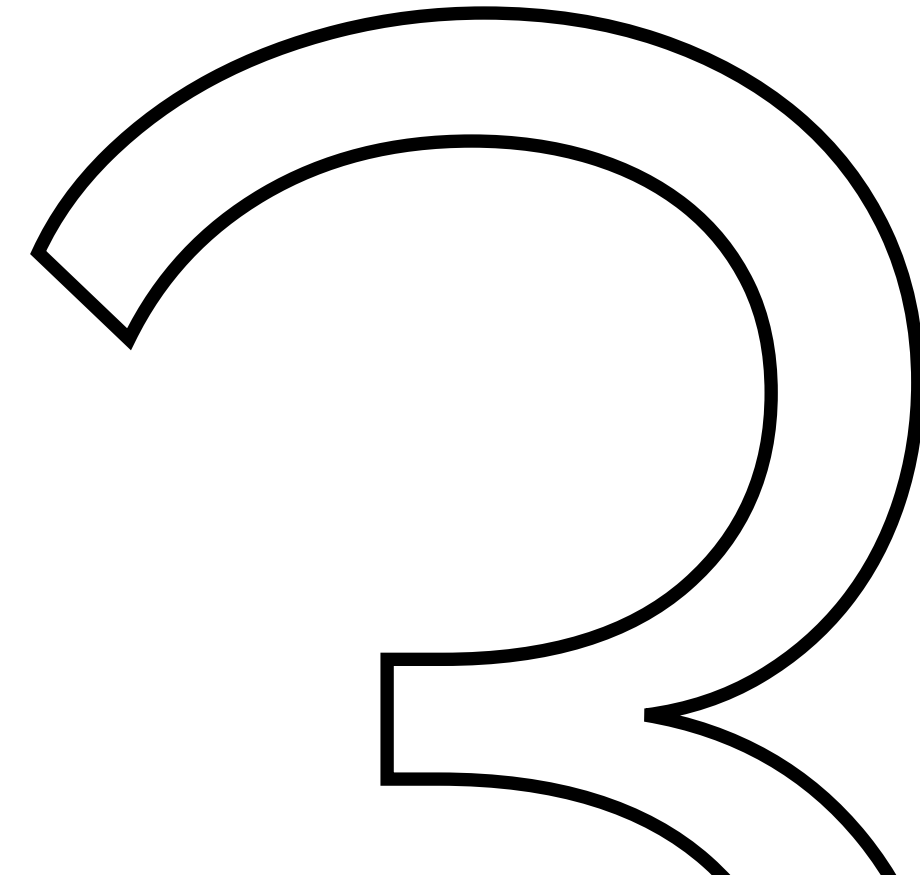
Elements of the academic literature on tribalistic language conformity and Liu's 2005 study on how people read online point to the importance of keyword identification and tribal signaling (Liu, 2005). Given this it's possible that the simple keyword substitution could actually be just enough to cause moderate dips in accuracy in responders. Also relying again on the understanding of the importance of language as a method of distinguishing fellow tribal members and opposing tribal members, it is also likely that responders that self identify as either Liberal or Conservative will be more susceptible to the effects of utilising these keywords in translations. If this is the case, the it is also likely that these individuals will outperform their non-tribalistic counterparts in the correct identification of the ideological origin of untranslated statements. Liu's insights on how reading speed

affects information processing also suggests that participants reading speeds will be directly related to their level of accuracy with substitution based translations, given skim reader's tendency to use "keyword spotting" to comprehend sentences (Liu, 2005).

This relationship between reading speed and accuracy will likely be less pronounced in survey 2's syntax and semantics based translations, as the changes are more subtle and may not be apparent to very quick readers. Overall, given the less obuse nature of Survey 2's translations, a more neutered or less extreme variation of the first survey's observations could result.

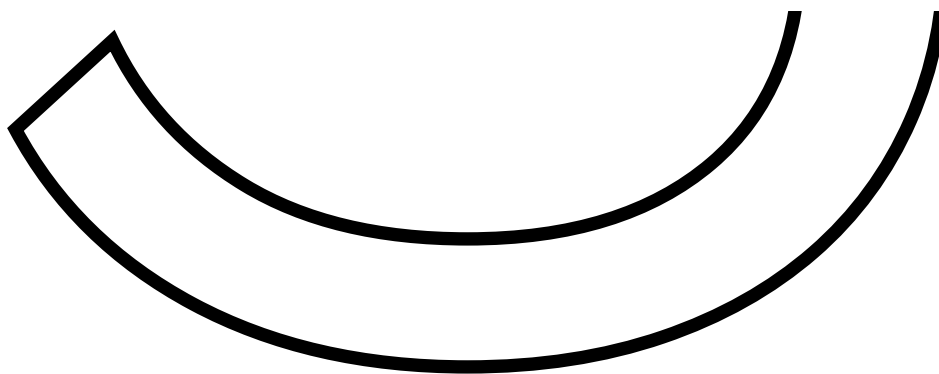
Survey 3 - Machine Learning

Machine learning translations, unlike substitution based translations may suffer from the implementation of the style transfer technology, and as a result yield less consistent results. However, given the grammatical complexity of translations achieved in the study by the creators of the back translation style transfer tool, successful outputs from the tool may well outperform all other translation methods. Less successful outputs, while equally valuable for testing purposes, may suffer from a decrease in sentence structure, resulting in less reliable data, or statements that will elicit a larger proportion of "Unknown" responses from participants. Overall, machine learning outputs may be a 'mixed bag' of some of the most successful translations, and the less successful.



C H A P T E R T H R E E

RESULTS



3.1 – RESULTS

Introduction

Running the surveys was a smooth process of utilising the specialised survey tool to upload new statements for testing from one of the outputs, removing old statements, and disseminating that survey to a new batch of responders. Speaking generally, many of the results observed fit nicely with previous understanding of how tribalism and language intersect. However, other elements of the survey were unexpected, especially the third and final survey's results.

Survey 1 Results

Responders

With 82 responders, survey 1 boasted the highest responder count among the surveys. It also provided the clearest window into the possible effectiveness of this process of translation. Of the 82 responders, 32 identified as liberal, 27 conservative, 16 neutral and 5 unknown. Two further responders were disqualified from the results for not finishing the survey, and having clearly unreasonable response times between statements, with a full survey completion time of 1 hour 45 minutes. Leaving us with 80 usable participants (fig 1).

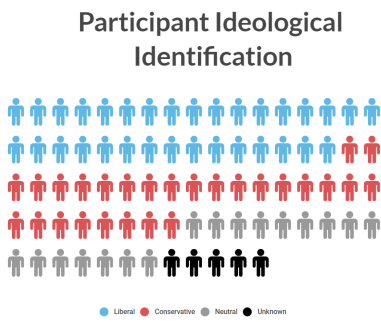
General Accuracy

Taking each responder's average response time and average accuracy rating (fig 2), we can observe that the longer any observer took to respond to a statement was correlated with a higher level of accuracy in identifying untranslated statement's political leaning, and a higher level of accuracy when identifying translated statement's original political leaning. However, these longer reading times were far less prevalent than the quicker alternative.

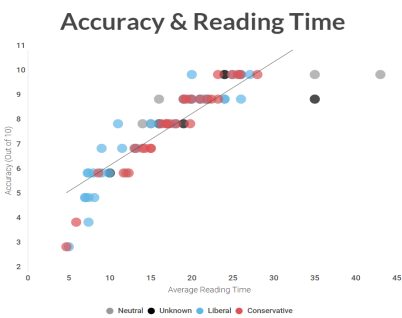
Dividing responders by their self assessed ideological leaning, we can see differences in average time spent reading (fig 3), and average accuracy of responders (fig 4). Neutral and Unknown affiliated individuals read statements longer and boasted higher accuracy ratings than their ideological counterparts. Within the Liberal and Conservative responders, Conservatives read for slightly longer and boasted a marginally higher accuracy rating than that of liberal responders. However this difference is very slight considering the small pool of responders.

Translated vs Untranslated Accuracy

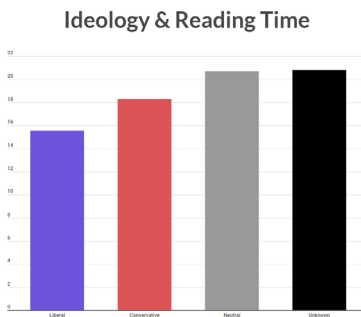
All responders were highly accurate at discerning the political leanings of untranslated statements, with responders averaging 95.2% (381 out of 400) accuracy in this field. Politically affiliated responders performed better



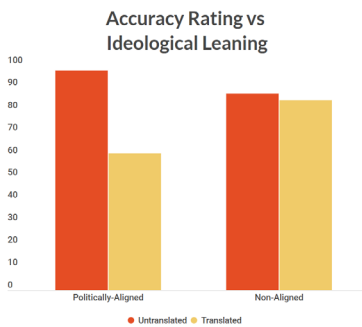
[Fig 1] Participant Ideological Identification



[Fig 2] Accuracy Rating vs Reading Time



[Fig 3] Survey 1 Reading Time vs Ideological Leaning



[Fig 4] Survey 1 Accuracy Rating vs Ideological Leaning

than their non-political comrades, with an average of 97.9% accuracy, with only 6 statements out of a collective 295 being mis-categorised compared to unaffiliated responder's 87.6%, with 13 mis-categorised out of 105 unchanged statements. This paradigm shifted heavily in the area of translated statements. Here, politically affiliated responders performed substantially worse than neutral or unknown responders. Politically unaffiliated responder's accuracy rates only shift marginally with translated statements, ending on 84.7%, whereas Liberal and Conservative responders showed a huge decrease in accuracy, down to 61% accuracy. Within this group, Conservatives came out on top by a razor thin margin, despite longer reading times on average. Reading times were also affected by the translations. Across the board reading times decreased when dealing with a translated statement as compared to untranslated. This decrease was more prevalent with politically affiliated responders, and within that group Liberals showed the largest decrease in reading times for translated statements.

Statement Specific Analysis

The translated statement that caused the lowest level of accuracy within Conservative groups was "Lying Hillary actually promises more jobs," where only 45% of Conservative responders identified the statement based on its original liberal sentiment. For Liberals, the translated statement that caused the lowest accuracy rating was "Shambling White House Prepares Executive Order to Weaken American Citizens Access to Healthcare." With a similarly low 43% of individuals identifying the Conservative origin of the statement. In both politically aligned groups, the statements with the highest success rates were the statements that had been translated to their ideology, from the opposing ideology. This trend was not visible in Neutral or Unknown responders.

Conclusion

Overall, political affiliation and time spent reading each statement were the prime predictors for accuracy. With time spent reading consistently providing a boost in accuracy and political affiliation providing greater accuracy with untranslated statements, but significantly lower accuracy with translated statements. Lacking a clear political alignment seems to offer less extreme highs and lows of accuracy, instead showing a more consistent accuracy rating. It also seemed that translated statements reduced the time spent reading by responders, possibly adding to the likelihood of inaccuracy.

Survey 2 Results

Responders

With a pool of 79 responders, the second survey retained a similar ideological makeup as the first. With 30 Liberal responders, 24 Conservative,

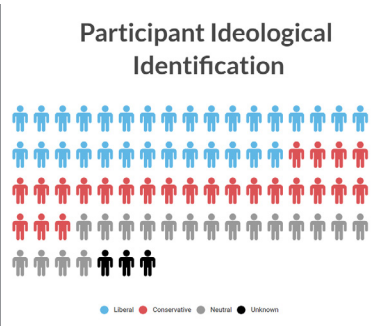
18 Neutral and 3 Unknown responders. However 4 responders needed to be disqualified for not finishing the survey, unreasonably long response times and unreasonably short response times, with one respondent completing the 10 statement survey in under 5 seconds. Leaving a total of 75 responders (fig 5). Survey 2 tested slight semantic and syntax changes to sentences and the effects that these changes had.

General Accuracy

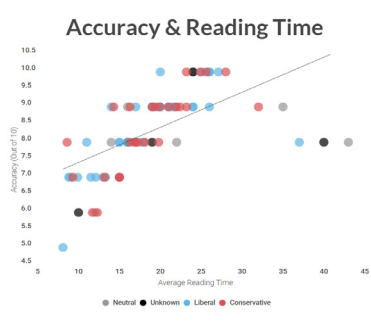
Overall this survey showed a less extreme variation of the results seen in the previous survey. Accuracy, on average showed a moderate increase across all responders. A similar division of reading times presents itself, with Neutral and Unknow responders taking longer to read the statements than politically affiliated participants. However, a key difference between this survey and the previous is the looser relationship between time spent reading and accuracy. The previous survey showed a consistent rise in average accuracy relating to a rise in the average seconds spent reading by participants. While this trend still shows in this survey, it is much less clear, and even begins to show that participants with the highest average reading times actually begin to decrease in accuracy, creating a very slight bell curve in the time spent reading and accuracy relationship (fig 6). Another key difference is that Neutral and Unknown responders perform worse on average than Liberal and Conservative participants (fig 7).

Translated vs Untranslated Accuracy

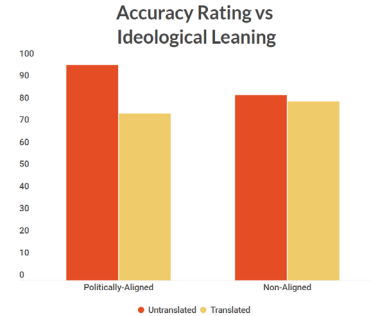
Accuracy in terms of identifying unchanged statements remained high across both politically affiliated and non-politically affiliated groups, though slightly lower than the previous survey with 93.6% accuracy (351 out of 375). A different relationship presents itself between responder’s political alignment and their accuracy when responding to untranslated statements than seen in the previous survey. With unaffiliated participants falling short of their politically aligned counterparts by a larger margin than the previous survey, with 83.8% (88 out of 105) compared to politically aligned responder’s 97.4% (263 out of 270), despite longer average reading times. Similar to survey 1, this trend of accuracy changed heavily when dealing with translated statements. However the change was not as severe as the first survey, a familiar pattern appears. Unaligned responders dip in accuracy slightly, down to 80.9% (85 out of 105), a significantly smaller drop than the previous survey. Similarly, Liberal and Conservative participant’s accuracy dip was much less intense, with a 22% drop to 75.5% (204 out of 270) compared to a 33% drop in accuracy witnessed in the first survey. So again, unaligned responders retain a higher accuracy rating than their politically aligned counterparts when responding to translated statements. However, in this case they fail to benefit from the extreme dip in accuracy



[Fig 5] Survey 2 Responder Ideological Identification



[Fig 6] Survey 2 Reading Time vs Accuracy



[Fig 7] Survey 2 Accuracy Rating vs Ideological Leaning

seen in the first survey, and end up with lower average accuracy (82.3%) than politically aligned participants (86.4%).

Statement Specific Analysis

The most effective forms of translation in this survey took the form of adding quote marks, most notably the phrase "'President' Trump strikes hard line against north korea" with the lowest accuracy scores across all participant groups (68%). Unlike the previous survey, there is no clear correlation between a responder's ideology and the types of translated statements more likely to result in further inaccuracy.

Conclusion

Overall the smaller changes yielded smaller results and some of the trends present in the first survey did not present here, most notably, the absence of a consistent time reading/accuracy relationship. The upper ends of reading time actually tended to lean slightly towards inaccuracy compared to responders with more moderate reading times. The non-aligned responder group's more consistent, approximately 80% accuracy ratings remained, but the more extreme dips in accuracy when dealing with translated statements did not present themselves in politically aligned groups. Leading to another of the key differences, that politically aligned responders actually performed better with both translated and untranslated statements. Translations such as de-capitalisation of proper nouns and important terms showed the least change in accuracy of responders, whereas the addition of quotation marks in key locations did create a drop in accuracy, though more generally, and not ideologically specific.

Survey 3 Results

Responders

Survey 3 had the lowest response rate of all three surveys, likely due to being stopped only after a short few days online. With only 12 responders, the sample size is likely too small to gather much information, however there were still interesting results within this group.

Conclusion

The machine learning algorithm produced translations that were difficult to understand, and that changed the meaning of the original statement to a much greater degree than the two previous surveys. The translations created were disjointed and mostly un-readable.

The survey was conducted nonetheless, but pulled after every responder identified each translated statement as "Unknown." Response times were shorter on average than previous surveys, and the data for the untranslated versions

remained more or less consistent with previous statistics despite the tiny sample size. Overall the data from this survey is mostly unusable, other than as a way of critiquing the potential of this particular machine learning translation technique.

ORIGINAL	TRANSLATED
Mr. McClintock - Please don't let Hillary get away with the lies she told to the American people about Benghazi.	Mr. Franken - Thank you for standing up to not attending Hillary Clinton has been voting against the way that it's going to the

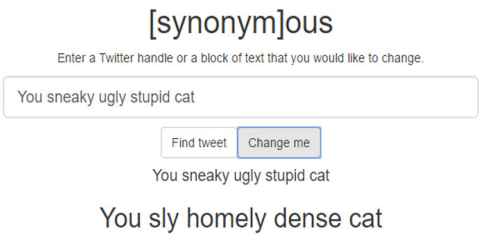
[Fig 8 - Left] Example Machine Learning Translation

Additional Sentiment Translation Tool Survey Results

With only 38 responders, this additional survey may not hold the keys to true understanding for this topic, however the results do illicit interesting conclusions regarding the possible effectiveness of simple word substitution in changing people's impressions of a statement. A key difference between this survey and the previous surveys is that this survey asks responders to judge a statements sentiment, not identify what ideology a sentence belongs to or is trying to propagate. When a statement is converted from positive to negative, while some of the original meaning remains, the sentence sentiment has changed, whereas in the political translations, most sentences continue to mean what they had previously meant. For example, the phrase “Lying Hillary promises more jobs” still promotes Hillary as a job creator. However the phrase “You sneaky ugly stupid cat” is negative in sentiment, regardless of the fact that the original sentence was not that negative. With that in mind, the term “accuracy” will continue to mean the act of identifying the sentiment of the original version of a translated phrase, or the act of correctly identifying an untranslated phrase. This is just to remain consistent between surveys for simplicity's sake.

General Accuracy

Similar to previous surveys, there is a relationship between time spent reading and accuracy. However the difference here is that this relationship only remains consistent to a point. Once responder's time spent reading reaches a certain level, it seems to stop having as powerful an effect as an accuracy predictor. Most of the inaccurate identifications of untranslated statements, and most of the 'accurate' identifications of translated statements occurred when responders took very little time to read the statements. Once a participant's



[Fig 9] Example Sentiment Translation from the '[Synonym]ous' Tool.

average time spent reading reached a tipping point, their accuracy in identifying untranslated statements and accuracy in identifying translated statements became rare.

Untranslated vs Translated Accuracy

When judging unchanged control statements, respondents were reasonably adept at discerning the sentiment of a given statement, with a 85.2% success rate. However the accuracy rate falls heavily when dealing with the translated statements. Statements translated from negative sentiment to positive sentiment produced the lowest accuracy ratings with just 18.9% (18 out of 95) of responders identifying the phrase as negative. With phrases translated inversely, the accuracy was slightly higher at 24.2% (23 out of 95), but remained significantly lower than anything seen in the previous, politically charged surveys.

Conclusion

This particular translation tool seemed to be so effective in translating a sentence sentiment that participants identified translated statements with the sentiment the statement had been translated to at a comparable rate (78.4%) to identifying unchanged statements accurately (85.2%). Time spent reading vs accuracy again created a different relationship, though followed the basic format of an increase in time spent reading leads to a higher chance of accuracy.

Results Overview

While survey 3 suffered a large setback in the difficulty of working with the machine learning algorithm, all surveys, including that setback gave great insight into the potential effectiveness of ideological translation tools. Each translation tool decreased accuracy effectively, although to different degrees, each tool predominantly affected members of the tribalistic ideological groupings and each tool showed a relationship (although not a consistent one) with time spent reading/attention spans and respondents accuracy. With larger word substitutions (Survey 1), the change in accuracy caused by a respondent being presented with a translated statement rather than a original statement was a drop in accuracy of 28% from 95.2% down to 67.2%. Across the two effective political base surveys, an average accuracy drop of 22% presents itself. From 94.4% accuracy in identifying statements that have not been changed, vs 72% accuracy with translated statements. Survey 2 certainly provided higher accuracy across the board, but overall the translation techniques experimented with here were successful with nearly 30% of respondents, even higher with those that the tools aim to target, the politically active and aligned.

4

C H A P T E R F O U R

DISCUSSION

1

4.1 – DISCUSSION

Introduction

Frequently, as designers, we are tasked with being on the forefront of providing avenues for positive change. Through sustainable design we can make positive impacts on mankind's environmental shortcomings, through interaction design we can make technology of all kinds more accessible to a larger variety of individuals, from physically disabled people through to the mentally challenged. Without a designer thinking outside the box we would never have experienced the joy of a fidget spinner. The role of design, as seen by some, is to create solutions to problems societies face. Political discourse in 2018-2019 has been marred with issues such as the problematic use of technology to perpetuate tribalistic tendencies and as a selective exposure based response to cognitive dissonance. It may be time for designers to take it upon themselves to explore the nature of tribalism and its negatives, to discover new ways of undermining or deconstructing the tribalistic nature of human society, with a goal of facilitating discourse that avoids entrenchment and hostility. The three design outputs tested here represent a step into the realm of searching for potential solutions to our destructive discourse and tribalistic tendencies. The results of these surveys show that translation tools have the potential to use the entrenched tribalistic language norms to the advantage of breaking through our selective exposure to information, while also shedding light on how future developments may build upon these findings with greater success. Surveys 1 and 2 both reflect the theories of modern political tribalism discussed earlier, while also showing a partial success in enacting the intended effect of tricking the reader into misidentifying statements by using simple substitution methods. However survey 3 shows that the realm of machine learning translation still has room to grow and improve before it can be properly applied to this design proposal.

Tribalism & Language Norms Presenting in the Results

The results of testing the design outputs show a reflection of the existing literature of political tribalism and linguistic study, while also showing the power of utilising tribal dialects as a way of combating cognitive dissonance among tribal participants. The language norms used in the first two surveys were intrinsically linked to the language norms and dialects of politically engaged and aligned individuals, as all of the statements were taken from political discourse. Politically aligned responders were more likely to recognise and correctly identify the ideologies of unchanged/untranslated statements. Their knowledge of the language norms, dialects and communication patterns due to their tribal alignment could have helped them gain that higher average accuracy among untranslated statements. Similar to the original examples of the tribes of Papua New Guinea and cliques of the American high school, participants involved in tribal groups

benefited from a heightened sensitivity to the use of language and performative norms. Unaligned participant's accuracy results showcased that the use of language norms associated with ideologies did not have as extreme an impact on effecting the responders answers. Whereas the politically aligned groups gained heightened accuracy with untranslated statements, and conversely, a larger drop in accuracy with translated statements. This initial observation shows a relationship between the insights of the results and the literature discussed surrounding the importance of the role tribalism plays in how an individual discerns the ideological origin of a written statement. In this particular case, it seems to be by focusing on the language associations relating to key words or phrases, rather than an holistic analysis of the full statement. So while a tribal alignment was related to higher accuracy with untouched statements, tribal affiliation was also linked to a greater drop in accuracy with translated statements. This could indicate that a key factor in a politically aligned participant's judgement of a statement is the how the statement conforms to the participants understanding of tribalistic language performance, in this case demonstrated by the use of key phrases and words.

This concept is similar to Sherif, Milroy and Margrain another of Milroy's findings on how language serves as a foundation to facilitate acceptance within a tribe, and as a way of identifying and recognising difference in intergroup discourse (Sherif, 1958), (Milroy, 1982), (Milroy & Margrain, 1980). These results also conform to Knobloch-Westerwick, Westerwick and Johnson's study showing individuals that "attach high importance" to the topics, tribalistic selective exposure tendencies are more prevalent. With a greater capacity to sort "attitude-consistent" material from "attitude-discrepant" material (Knobloch-Westerwick, Westerwick & Johnson, 2015, p 1). Viewing the results through this lens, it seems that when responding to untranslated statements, participants with a tribal ideology used their ability to recognise the key phrases, terms or vernacular they use to "facilitate acceptance" within their tribe to determine if a statement fit within their own ideology. These individuals also use the ability to recognise the language their tribe identifies as different or opposing to determine if a statement was of a differing tribal/ideological background. In effect, showing a capacity to sort "attitude-consistent" material from "attitude-discrepant" material. This recognition and categorisation based the presence of tribal terminology could also be what in turn led to the greater decrease in accuracy seen within politically affiliated participants in comparison to unaffiliated respondents.

This decrease in accuracy is what this research had predicted based on the available literature. With such simplistic language changes providing relatively large drops in participants ability to correctly define a statement's political leaning, these results show a possibly effective way of mitigating tribalistic online media user's proclivity to omit opposing ideological messaging. If the ideology within say, a headline or title of an article, is harder to immediately identify as an opposing one, it is possible that a reader may not be so hasty to disregard the article on the

basis of combatting their own cognitive dissonance.

Reading Time & Accuracy

Another important relationship to observe within these results is the relationship between reading time and accuracy. With an average drop in accuracy of 22% when assessing translated statements compared to untranslated statements, these linguistically basic substitutions of key phrases performed relatively well. However, the lower accuracy was heavily linked to a low amount of time spent reading each message. This indicates that at least with substitution based methods, it may be difficult to successfully lower the accuracy of people who spend a higher amount of time reading online messaging. The longer an individual, political or otherwise, reads a statement, the more they are likely to have a greater comprehension of the overall meaning of the text. In a 2002 study, Dyson and Haselgrove discovered a speed vs accuracy trade off in modern online text consumption. Specifically within sites where larger amounts of scrolling were required (Dyson & Haselgrove, 2002). Related too is Dymock's discussion of the 'word caller' or an individual who reads text by monitoring each word individually, rather than the sentence as a whole (Dymock, 1993).. Dyson & Haselgrove's discussions on the increase in skim reading in the internet age is particularly relevant here, as it is unlikely that responders to this survey were reading at the same speed at which they regularly browse the internet. It is likely that given that the participants knew they were performing a survey, their reading speeds would have altered, potentially lengthened given the survey environment compared to the scrolling environment of a modern news site, message board or social media platform. This could indicate that if the translated statements had somehow been exposed to individuals within these scrolling heavy environments, the accuracy drop could have been even greater as readers would have returned to their more casual reading speeds.

For readers with high reading times even on scrolling based platforms, it is likely that a more linguistically complex method of translation would need to be utilised to create a sizeable drop in that demographic's accuracy levels. Interestingly though, Survey 2 saw a drop in accuracy in the participants with the highest reading times. This could be an anomaly, or it could potentially represent an existence of a very slight bell curve within the results of the syntax based translation tests. This could be an indication that given greater amounts of time to read a statement, a participant can begin to search for meaning in the syntax changes of translated statements. Possibly leading them to read further "in" to the statement. If this is the case, this type of subtle translations could be used to target readers with higher screen reading times. However, given the small sample size of this survey, it is certainly not a conclusive element of this study. Other than that accuracy drop in the higher reading rates, and a moderate shift towards higher accuracy overall, the trend of higher reading times relating to higher

accuracy ratings remained mostly consistent with Survey 1 and the findings of Dyson and Haselgrove. Overall, it seems that online media consumers provide a perfect audience for these translation tools, given the relatively low reading speed and accuracy associated with this demographic.

Selective Exposure, Cognitive Dissonance & Skim Reading

Interestingly, if modern internet users tend towards skim reading in scrolling environments, that also asserts that within this text consumption method, any cognitive dissonance created by reading an opposing ideology, or an uncomfortable message is generated from this quick and low accuracy reading style. This implies a reliance on quick methods of categorising written phrases, to match the speed of reading. The results of the first two surveys could provide some insight into the types of key phrases that determine if cognitive dissonance is generated and in turn if an individual decides to avoid that information on the basis of selective exposure. Within the two surveys, politically aligned respondents were, on average, more likely to read translated statements for a shorter amount of time than untranslated statements. This indicates that whatever key words or phrases that were substituted into those translated statements helped those responders to make their decisions quicker, and in turn, generally less accurately. This implies that statements read naturally in online environments that use ideologically charged phrases or words are more likely to elicit a response of either tribalistic familiarity or create cognitive dissonance at a quicker rate than a more neutral statement, and to a degree that may impact an individual's ability to correctly assess what ideology the statement truly holds. As discussed in Liu's article on reading behaviour in the digital age and the changes in reading behaviour over the past decade, modern digital readers are more likely to rely on scanning, and keyword spotting (Liu, 2005). Liu also discusses a decrease in in-depth reading and an increase in reading selectively, or using selective exposure to determine the texts that are read, based on the results of the aforementioned habits of scanning and keyword spotting.

With this in mind, the results of the digital output tests become quite unsurprising. A heavier reliance on scanning and keyword spotting would explain the large drop in accuracy seen with translated statements, as participants relying on keyword changes for their judgements would be more likely to categorise translated statements based on the tactically chosen keywords. Importantly, if an individual's media consumption is selective based on their use of scanning and keyword identification, any method of successfully substituting keywords into statements has the potential to alter how that individual consumes media.

Applications and Rationale

Academic discussion of modern news media and political division is underpinned by repeated references to the adverse effects of selective exposure and manifestations of tribalisation such as filter bubbles or echo chambers (Bunch of References). This thesis intends to utilise the vast resources of political, sociological, psychological and linguistic study, combined with elements of natural language processing and machine learning techniques to show a proof of concept for design solutions to target people's tribalistic, and selective exposure based responses to experiencing dissonance when consuming divisive media, in this case, political news and discourse. The possible connotations of the results discussed above provide a window into the potential success of a more cohesive, streamlined and precise version of the design outputs tested here. In these design outputs, even the most simplistic linguistic changes still caused relatively sizeable drops in accuracy, enough to serve as the basis for further research, design iteration and testing. Westerwick, Westerwick and Johnson's study showed that individuals that placed higher value on issues were more likely to identify and selectively expose themselves to attitude-consistent media. They also showed that while individuals involved in their studies did show preference for high credibility media, this factor did not mitigate the effects of their selective exposure. Meaning that low credibility attitude-consistent material was often favoured over high credibility attitude-discrepant material. Finally, their study concluded that these factors also led to shifts in attitudes among participants (Knobloch-Westerwick, Westerwick & Johnson, 2015). These observations allude to the power of modern discourse and information technology, and how participants in political discourse can be likely to fall into the habits of tribalistic tendencies such as selective exposure and echochambers. The results from both survey 1 and 2 show findings consistent with elements of these studies. It was also observed that responders with ideological alignments were similarly adept at identifying attitude-consistent and discrepant dialects/terminologies. These overlapping findings, and the results showing dips in accuracy following translations in survey 1 and 2 suggest that these design methods could offer avenues to combat politically aligned individuals' heightened ability to identify consistent/discrepant terminologies, by using their heightened and sensitive skim reading, keyword based identification method against them. By disguising information from attitude-discrepant sources in the tribalistic language norms of attitude-consistent information.

Applications of this type of design tool could be broad, from advertising, to conflict mediation and through to political campaigning. While this tool has been tested as a means to bridge political divides, the implications for a resource that has the potential to lessen the effects of cognitive dissonance reach broader than the realm of political activism. However, it is important to note that the intention behind this research has been to suggest methods to help slow the growth in political partisanship, rather than as a utility for general information

dissemination and sales. The intended applications for this research would instead sit firmly within the realm of tools used to re-present news of a greater variety and scope to audiences in the hopes of creating a more positive political discourse

Concerns & Moving Forward

environment.

Despite successes in the testing of the design outputs for this research, the process encountered several insights due to failures, and elements of the survey design should be revised and improved upon for future studies. A machine learning based translation method was the originally planned final iteration of this project. The prediction being that results would show the improvements in translation success between the substitution methods and more advanced algorithmic solutions. However that was not the result. The technology utilised was a machine learning style transfer tool, created through Pittsburgh University. This particular tool was chosen due to its successful implementation in Prabhumoye, Tsvetkov, Salakhutdinov and Black's 2018 study on back translation and its free, open source availability. Within this implementation, the tool showed success in translating political statements from conservative to liberal and vice versa (Prabhumoye et

[Right] Example translations of machine learning tool from Pittsburg back translation tool (Prabhumoye et al., 2018).

[Below] Example translations of machine learning tool implimented for these surveys.

ORIGINAL	TRANSLATED
Mr. McClintock - Please don't let Hillary get away with the lies she told to the American people about Benghazi.	Mr. Franken - Thank you for standing up to not attending Hillary Clinton has been voting against the way that it's going to the

republican → democratic		
<i>i will continue praying for you and the decisions made by our government!</i>	<i>i am proud of you and your vote for us!</i>	<i>i will continue to fight for you and the rest of our democracy!</i>
<i>tom, i wish u would bring change.</i>	<i>i agree, senator warren and could be.</i>	<i>brian, i am proud to have you representing me.</i>
<i>all talk and no action-why dont you have some guts like breitbart</i>	<i>and then we will be praying for them and i am proud of this position and i am proud of</i>	<i>keep up and dont know, you have a lot of respect as breitbart</i>
democratic → republican		
<i>as a hoosier, i thank you, rep. visclosky.</i>	<i>a lot, i am proud of you <unk>.</i>	<i>as a hoosier, i'm praying for you sir.</i>
<i>thank you for standing up for justice and against bigotry-racism, homophobia, sexism, misogyny, religious and xenophobia.</i>	<i>do you for standing up for highly and in bigotry-racism, programming, cut, granddaughters, unprecedented and excludes.</i>	<i>thanks for standing up for the constitution and get rid of obamacare, homophobia, cut, and actuality.</i>
<i>thank you for all you are doing for us, attorney general harris!</i>	<i>thank you for standing up for us and i am proud of us!</i>	<i>thanks lawmaker for all you do for us, senator scott!</i>

al., 2018).

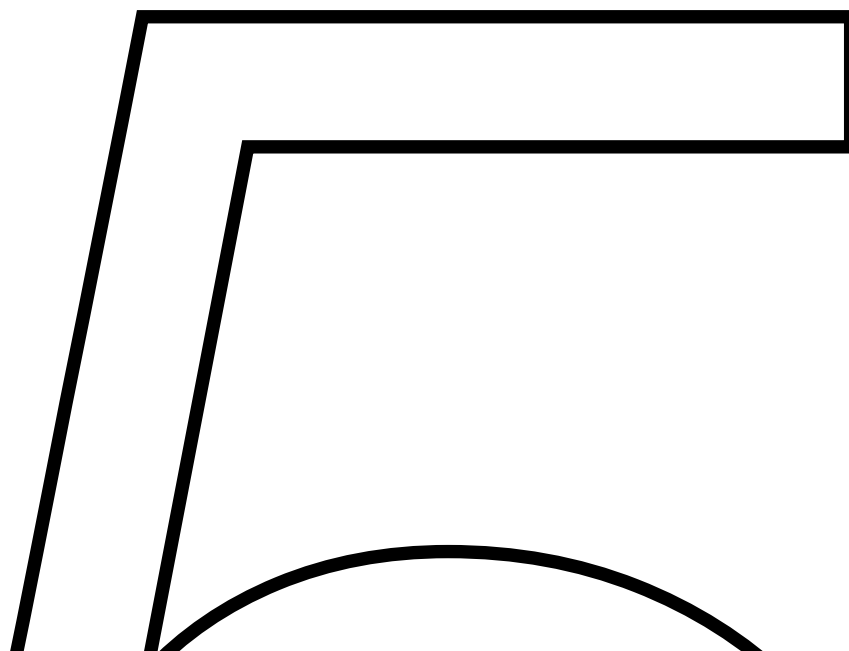
In this project's application of the tool, however, results such as these were not replicated, though instructions provided with the tool were followed diligently and accurately. Instead, the resulting phrases were nonsensical and difficult to penetrate. Through brief testing discussed in the results section, it was determined that participants simply couldn't understand the phrases enough for any decent information to be gathered and the test was halted.

Despite this setback, this process has outlined possible limitations of this

particular style of transfer tools, assuming that there was not an undiscovered error undermining these outputs. In their research, Prabhumoye, Tsvetkov, Salakhutdinov & Black showed that the Style-Transfer tool used in this thesis had consistently outperformed other translation/transfer methods in terms of style accuracy and meaning preservation. However, despite outperforming other methods, this tool's meaning preservation performed at a mediocre level when tested by human evaluation methods (Prabhumoye et al., 2018). This, combined with a redefinition of the term "meaning preservation" to be more broad, that while reasonable contextually within the original study, did not meet the standards for meaning preservation required by this research. Given that this 2018 variation of a style transfer tool outperformed its 2017 counterparts by over 30% in accuracy and meaning preservation, it is reasonable to believe that this technology will continue to grow quickly, and will soon be more than capable of performing the translations required for the purpose of this research (Prabhumoye et al., 2018). However, in this study, the machine learning implementation output severely underperformed compared to its substitution based precursor iterations.

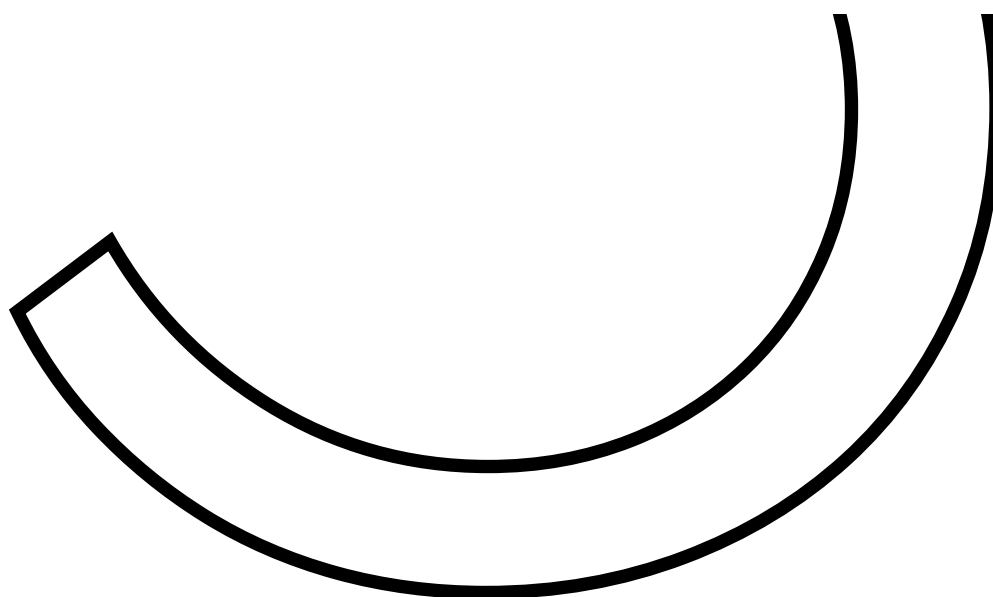
Other elements that could have benefitted this study include; a more nuanced way for participants to politically identify themselves than multi-choice options including Liberal and Conservative, including a question within the survey that could judge a participant's political engagement level and an extension of the research to measure how much sense participants could make out of translated statements. If a spectrum of political alignments was offered, rather than the classic binary political party structure, it could have been possible to measure if an increased subscription to a tribalistic ideology caused further dips in accuracy within responses. The current surveys successfully showed a correlation between identifying politically and being affected by keyword translations, however from the data recovered it is not possible to delve deeper into this result for further insight. Equally, as seen in Westerwick, Westerwick & Johnson's 2014 study, it would have been equally valuable to also measure respondents self identified levels of engagement in political matters. The largest recommended addition to future research in this area would be to extend the scope of the survey's to also measure the meanings responders interpret from the translated statements. As it would be beneficial to understand whether the respondent absorbs elements of the translated statement's original ideology or if they convert the whole statement to be attitude-consistent.

Overall the results discussed here point to the potential of digital design solutions to linguistic and discourse based problems, however with further and deeper evaluation, the concept could be expanded beyond potential into functioning methods of political activism.



C H A P T E R F I V E

CONCLUSION



5.1 - CONCLUSION

Some elements of the academic discussion of 'tribalism' attaches a dark and barbaric connotation to the term, described as primitive, or archaic, or something that we must rise above at all costs (Smith, 1981) (Hutcheon, 2001). While these are fair critiques, tribalism has been a hugely positive influence on our societies for ages (Plater, 1990). The natural urge to group together and find common goals, attitudes, likes and dislikes has helped shepherd us through hostile environments and continues to fill the world with a huge array of vastly different cultures, subcultures, trends and fads (Kysar and Salzman, 2003) (James, 2006). These shifting, morphing elements of conformity and cultural groupings are natural representations of our innate tribalistic urges, and that does deserve some celebration. It seems that wolves hunt in packs, birds of a feather, flock together and maybe humans exist in tribes. However, tribalism can be a blunt instrument when it comes to creating social formations. The desire to belong to something can inherently come at the expense of excluding, or purposefully creating divides between others. The concept of "we are us, because we are not them" arises, which can then shift towards "we can only remain us, if we obliterate them" (Volf, 1998). This dogmatism has been visible in a host of different tribalistic structures throughout history and continues today, through religious conflicts, xenophobia and elitism (Volf, 2002). It is this natural yet corrupted extension of tribalism that leads to the cognitive dissonance visible in the political discourse ecosystem. This dissonance is ultimately the root of the selective exposure, bias, filter-bubbles, echochambers and partisan divide witnessed in modern political media. Each of these social structures and mechanisms created via this dissonance contribute to growing divisions within modern political discourse, and while these occurrences may be the extension of a natural evolutionary urge, that does not mean they should be left unchecked.

This growing threat to the discourse that democratic practises rely on has been heavily investigated and hotly debated, which has lead to a plethora of resources to draw on when designing the translation tools and testing structures of this thesis' output. Combining elements of linguistic study of language identity and codeswitching, with the concept of selective exposure and tribalistic conformity, and finally with how online media is consumed, I theorised a translation tool seemed to show potential as a method of disrupting the cognitive dissonance that permeates throughout political culture. This hypothesis proved to be true, as after an iterative design and testing process, the results suggest that translation based tools do show promise with subverting the tribalistic associations of key phrases. The testing process also discovered key links reflecting elements of the broader research around tribalism by showing that participants who identified themselves as politically aligned were significantly more susceptible, than their unaligned counterparts, to translated statements that utilised key phrases identified with either liberal or conservative

tribal groups. Overall, with the aid of a wide range of research disciplines, this thesis has produced a foundation of effectiveness for future studies to develop and iterate upon, with the goal of using this technology to broaden the range of information all individuals expose themselves to, to prevent ideological insulation as a result of tribalistic cognitive dissonance.

WORKS CITED

- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130-1132.
- Bernhardt, D., Krasa, S., & Polborn, M. (2008). Political polarization and the electoral effects of media bias. *Journal of Public Economics*, 92(5-6), 1092-1104.
- Brock, T. C., & Balloun, J. L. (1967). Behavioral receptivity to dissonant information. *Journal of personality and social psychology*, 6(4p1), 413.
- Brownstein, R. (2016, September 1). Voters don't split ticket anymore. *The Atlantic*. Retrieved from <https://www.theatlantic.com/politics/archive/2016/09/senate-trump/498287/>
- Brundidge, J. (2010). Encountering "difference" in the contemporary public sphere: The contribution of the Internet to the heterogeneity of political discussion networks. *Journal of Communication*, 60(4), 680-700.
- Bucholtz, M. (1999). "Why be normal?": Language and identity practices in a community of nerd girls. *Language in society*, 28(2), 203-223.
- Bucholtz, M., & Hall, K. (2004). Language and identity.
- Chandler, J. (2013). Faith-based policy: A litmus test for understanding contemporary America. *Lexington Books*.
- Conover, M., Ratkiewicz, J., Francisco, M. R., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. *Icwsn*, 133, 89-96.
- Conover, M. D., Gonçalves, B., Ratkiewicz, J., Flammini, A., & Menczer, F. (2011, October). Predicting the political alignment of twitter users. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference* (pp. 192-199). IEEE.
- Cotton, J. L. (1985). Cognitive dissonance in selective exposure. *Selective exposure to communication*, 11-33.
- Dymock, S. (1993). Reading but not understanding. *Journal of Reading*, 37(2), 86-91.

Dyson, M., & Haselgrove, M. (2000). The effects of reading speed and reading patterns on the understanding of text read from screen. *Journal of Research in Reading*, 23(2), 210-223.

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annu. Rev. Polit. Sci.*, 11, 563-588.

Gainous, J., & Wagner, K. M. (2013). Tweeting to power: The social media revolution in American politics. *Oxford University Press*.

Garrett, R. K., Carnahan, D., & Lynch, E. K. (2013). A turn toward avoidance? Selective exposure to online political information, 2004–2008. *Political Behavior*, 35(1), 113-134.

Graham, T., Jackson, D., & Broersma, M. (2016). New platform, old habits? Candidates' use of Twitter during the 2010 British and Dutch general election campaigns. *New media & society*, 18(5), 765-783.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin*, 135(4), 555-88.

Hawkins, S., Yudkin, D., Juan-Torres, M., Dixon, T. (2018). *Hidden tribes: A study of America's polarized landscape*. Retrieved from <https://www.moreincommon.com/hidden-tribes/>

Hobfoll, S. E. (2018). Tribalism: The evolutionary origins of fear politics.

Huang, C. (2011). Facebook and Twitter key to Arab Spring uprisings: report. *The National*. Retrieved from <http://www.l2f.inesc-id.pt/~fmmb/wiki/uploads/Work/misnis.ref01.pdf>

Hutcheon, P. D. (2001). Can humanism stem the rising tide of tribalism?. *Humanist in Canada*, 34(1), 6-11.

James, P. (2006). Globalism, nationalism, tribalism: Bringing theory back in. *Pine Forge Press*.

Kinney, D. A. (1993). From nerds to normals: The recovery of identity among adolescents from middle school to high school. *Sociology of Education*, 21-40.

Knobloch-Westerwick, S., Johnson, B. K., & Westerwick, A. (2014). Confirmation bias in online searches: Impacts of selective exposure before an election on political attitude strength and shifts. *Journal of Computer-Mediated Communication*, 20(2), 171-187.

Kobayashi, T., & Ikeda, K. I. (2009). Selective exposure in political web browsing: Empirical verification of 'cyber-balkanization' in Japan and the USA. *Information, Communication & Society*, 12(6), 929-953.

Kysar, D. A., & Salzman, J. (2002). Environmental tribalism. *Minn. L. Rev.*, 87, 1099.

Lance Bennett, W. (2012). The Personalization of Politics: Political Identity, Social Media, and Changing Patterns of Participation. *The ANNALS of the American Academy of Political and Social Science*, 644(1), 20-39.

Liu, Z. (2005). Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of documentation*, 61(6), 700-712.

Milroy, L. (1982). Language and group identity. *Journal of Multilingual & Multicultural Development*, 3(3), 207-216.

Milroy, L., & Margrain, S. (1980). Vernacular language loyalty and social network. *Language in Society*, 9(1), 43-70.

Plater, Z. J. (1990). Keynote Essay: A Modern Political Tribalism in Natural Resources Management. *Pub. Land L. Rev.*, 11, 1.

Prabhumoye, S., Tsvetkov, Y., Salakhutdinov, R., & Black, A. W. (2018). Style Transfer Through Back-Translation. *arXiv*, preprint arXiv:1804.09000.

Resnick, P., Garrett, R. K., Kriplean, T., Munson, S. A., & Stroud, N. J. (2013, February). Bursting your (filter) bubble: strategies for promoting diverse exposure. In *Proceedings of the 2013 conference on Computer supported cooperative work companion* (pp. 95-100). ACM.

Sears, D., & Freedman, J. (1967). Selective Exposure to Information: A Critical Review. *The Public Opinion Quarterly*, 31(2), 194-213. Retrieved from <http://www.jstor.org/stable/2747198>

Shashkevich, A. (2017, January 3). Can Empathy Bridge Political Divides? *Greater Good Magazine*. Retrieved from https://greatergood.berkeley.edu/article/item/can_empathy_bridge_political_divides

Sherif, M. (1958). Superordinate goals in the reduction of intergroup conflict. *American Journal of Sociology*, 63(4), 349-356.

Skoric, M. M., Zhu, Q., & Lin, J. H. T. (2018). What Predicts Selective Avoidance on Social Media? A Study of Political Unfriending in Hong Kong and Taiwan. *American Behavioral Scientist*, 62(8), 1097-1115.

Smith, A. D. (1981). War and ethnicity: The role of warfare in the formation, self-images and cohesion of ethnic communities. *Ethnic and Racial Studies*, 4(4), 375-397.

Stieglitz, S., & Dang-Xuan, L. (2012, January). Political communication and influence through microblogging--An empirical analysis of sentiment in Twitter messages and retweet behavior. In *System Science (HICSS), 2012 45th Hawaii International Conference* (pp. 3500-3509). IEEE.

Sussman, S. , Dent, C. W., Mestel-Rauch, J. , Johnson, C. A., Hansen, W. B. & Flay, B. R. (1988). Adolescent Non-smokers, Triers, and Regular Smokers' Estimates of Cigarette Smoking Prevalence: When Do Overestimations Occur and by Whom? *Journal of Applied Social Psychology*, 18: 537-551.

Valentino, N. A., Banks, A. J., Hutchings, V. L., & Davis, A. K. (2009). Selective exposure in the Internet age: The interaction between anxiety and information utility. *Political Psychology*, 30(4), 591-613.

Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An Identity-based model of political belief. *Trends in cognitive sciences*.

Volf, M. (1998). "The Trinity is our social program": The doctrine of the Trinity and the shape of social engagement. *Modern theology*, 14(3), 403-423.

Volf, M. (2002). Christianity and Violence. *Boardman Lectureship in Christian Ethics*. Retrieved from <https://repository.upenn.edu/boardman/2>

Wodak, R. (Ed.). (1989). Language, power and ideology: Studies in political discourse (Vol. 7). *John Benjamins Publishing Company*.

Woolard, K. A. (2004). Codeswitching. *A companion to linguistic anthropology*, 73-94.

Zillmann, D., & Bryant, J. (2013). *Selective exposure to communication*. Routledge.

US VS THEM

Codeswitching is the action of switching between languages to better impart information to a recipient. This thesis introduces a set of codeswitching translator tools as a method of disrupting the potentially damaging structures of tribal politics through the manipulation of ideologically specific language norms. We first investigate how tribalism and group identity impact our ability to participate in political discourse. Using this insight from a host of different research disciplines, we design an iterative testing environment for a variety of 'codeswitching' translators in order to see the impact of translations ranging in complexity from simple word and syntax substitution through to machine learning back-translation. Though back-translation was not found to be an effective technique, simple substitution methods provided a foundation of effectiveness and proof of concept among test participants, especially those that identified as politically aligned.