Bangor University

DOCTOR OF PHILOSOPHY

How-to select which language to use in a human disaster: The roles of language planning, (mutual) intelligibility & linguistic attitudes.

Glackin, Conor

Award date:
2022

Awarding institution:
Bangor University

[Link to publication](#)

How-to select which language to use in a human disaster:

The roles of language planning, (mutual) intelligibility & linguistic attitudes

Conor Glackin

A Thesis submitted in partial fulfilment of the requirement for the degree of Doctor of Philosophy in Linguistics at Bangor University

School Of Arts, Culture and Language

Bangor University

July 2022

Declaration

'Yr wyf drwy hyn yn datgan mai canlyniad fy ymchwil fy hun yw'r thesis hwn, ac eithrio lle nodir yn wahanol. Caiff ffynonellau eraill eu cydnabod gan droednodiadau yn rhoi cyfeiriadau eglur. Nid yw sylwedd y gwaith hwn wedi cael ei dderbyn o'r blaen ar gyfer unrhyw radd, ac nid yw'n cael ei gyflwyno ar yr un pryd mewn ymgeisiaeth am unrhyw radd oni bai ei fod, fel y cytunwyd gan y Brifysgol, am gymwysterau deuol cymeradwy.'

Rwy'n cadarnhau fy mod yn cyflwyno'r gwaith gyda chytundeb fy Ngrichwyliwr (Goruchwylwyr)'

'I hereby declare that this thesis is the results of my own investigations, except where otherwise stated. All other sources are acknowledged by bibliographic references. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless, as agreed by the University, for approved dual awards.'

I confirm that I am submitting the work with the agreement of my Supervisor(s)'

Acknowledgements

Firstly, I need to thank my supervisor Dr Marco Tamburelli, for his unyielding support throughout the project. Every time I felt dumb, lost, or overwhelmed I would be pulled back on track with his wise words of guidance and his humorous take on all things. Thank you, Marco, for keeping my spirits up through the years and believing in me, even when I stumbled or went down another rabbit-hole. I'll admit that you were right all along, that I would finish the beast that is this thesis.

Secondly, I need to thank every Arabic speaker who interacted with the project. My deepest gratitude goes to those who helped with producing and recording experimental stimulus. Special thanks go to Saher Abdulrahmman, without you this project would not have been possible, and I hope to work with you again in the future.

Thirdly, I need to thank my family and friends for their never-ending faith that I could get to the end of the process. This project relied on your collective support, whether that was a quick motivation message or a three-hour rant over social theory. Special thanks to Lucy Smallwood, for persevering through the drafts of this thesis and providing invaluable feedback with disaster response level haste.

My final thanks goes to my husband Ben. Without your continuous support, and unshakable faith in my abilities, I would not have obtained a PhD placement, yet alone manage to complete the thesis. You have been there for every euphoric high and each spiral of despair; and not once have you ever given up on me, or the project. I remember telling you that my life goal was to complete a PhD, and you casually dismissing my worries that it could never happen. Maybe, just maybe, your confidence was not misplaced.

Abstract

This thesis explores language policy in situations of human based disasters. This project is motivated from the call for action declared by a 2017 Translators Without Borders (TWB) report on the on-going Southern European refugee crisis; particularly the widespread communication issues that exist daily within the aid camps. Refugees are vulnerable individuals, who have experienced hardship, and are being processed without adequate language support; as such they have no home, no nation to protect them, and cannot be understood. As a result, refugees quickly become disenfranchised with the outstretched support networks, and begin to disregard aid advice and become susceptible to radicalisation. Even after language support and communication is established, distrust can still develop, as refugees use stereotypes to judge the aid workers, such as, if the aid worker uses language which is associated with a rival society, then negative underlying language attitudes result a low perception of veracity in the messaging.

This thesis argues that, at its core, these communication issues are a result of insufficient language planning for appropriate, accessible, and compliable policies for use in the active field; as well as being the repercussions of widely established issues in sociolinguistics and language demarcation. The prominent issue of what defines a language itself, is ever-present in the situations documented by TWB (2017). Kloss (1967) split language demarcation criteria into two ideological groups; language by either socio-political opinion (Ausbau) or by linguistic differences (Abstand). This was to allow for a dual classification of language, in a way to ensure that communication was measured alongside the political identities when declaring a

language. However, this dual approach was largely disregarded due to operational issues in Abstand. As a result, globally there is a default bias towards using Ausbau definitions, with a defacto status of Ausbaucentrism (Tamburelli, 2021), whereby individuals are grouped as being speakers of the same language, yet they cannot speak to each other.

In chapter three, this thesis explored the impact of Ausbaucentrism on international language policies, by proposing a method to objectively identify which languages should be used in the event of a disaster. This method, called Reach, focuses on the maximum impact of a language policy, specifically, whether the population impacted will be able to get to access information and other resources in a language they can understand, i.e., that the language policy is intelligible. From this, the issue of mutual intelligibility within the diglossic Arabic language was identified as a research gap. Chapter four explored this gap with an experimental study measuring the mutual intelligibility of three Mashriqi Arabic varieties (Nijari, Cairene and Gulf-Modern Standard); in a design which simulated parts of the socio-environmental context as those from the TWB report, namely, distress. Finding that mutual intelligibility in Arabic cannot be assumed; in fact, providing evidence for the splitting of Arabic into multiple languages. Chapter 5 investigated underlying distrust of a speaker based on language attitudes between the three Arabics. Finding that the prestige and familiarity of an Arabic variety does not infer that the speakers will be automatically trusted most. In fact, this study identified that the unfamiliar international neighbours are perceived more positively than the local neighbours.

Overall, this thesis provides three methods to quantify vital aspects needed to produce adequate language policies for use in disaster response. These methods, when combined, provide a detailed picture of whether a language should be used in disaster policies, when the three primary criteria are: accurate communication; accessibility of a language and Abstand. Thus, in part filling addressing the overarching research gap, of how to measure, and then counter-act widespread communication issues in disasters.

Contents

List of Tables

List of Figures

List of Appendices

Abbreviations

ACC	Accuracy Rate
AL	Arab League
AU/ OAU	African Union (previously call Organization of African Unity)
CA	Classical Arabic
CALD	Contents, Associations, Length and Delivery
CDC	Centre for Disease Control and Prevention
CLT	Cognitive Load Theory
CPT	Population of year of closest entry
CSN	Closest Year Entry
DRC	Democratic Republic of the Congo
EA	Egyptian Arabic
EGIDS	Expanded Graded Intergenerational Disruption Scale
EU	European Union

EULP	European Union Language Policy
GA	Gulf Arabic
GIDS	Graded Intergenerational Disruption Scale
H	High diglossia variety
HP	High Predictability
IELTS	International English Language Testing System
L	Low diglossic variety
LP	Low Predictability
MSA	Modern Standard Arabic
NGO	Non-Government Organisation
PT	Population in the Target Region
PTSD	Post-Traumatic Stress Disorder
PTY	Population in the Target Region of a Targeted Year
QA	Colloquial Arabic
RT	Reaction Time
SA	Standard Arabic
SCCT	Situational Crisis Communication Theory
SIT	Social Identity Theory
SMS	Short Messaging Service
SN	Speaker Numbers
SPIN	Speech Perception In Noise Test
TROG-2	Test for Reception Of Grammar-2
TWB	Translators Without Borders
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNGA	United Nations General Assembly
UNLP	United Nations Language Policy
WM	Working Memory

Chapter 1: Introduction

1.1 Considerations of the Peripheral Elements

When considering language policy and language planning, Wright (2016) highlights that the discipline reflects political, economic & social processes and events, as that to ignore the connection between language and those who use the language is a shortcoming of sociolinguistics. Within the study of historical linguistics and historical languages, it is widely acknowledged that noteworthy events across history have [in the timeline] affected the language practices of whole countries (Weiss, 2015). It is difficult to see how historic events, and any subsequent changes in language practices occurred if not for a form of language policy or practice intervention by organised bodies or social groups. Ó Riagáin (1997:170-171) argued that "language policies to produce intended outcomes is severely constrained by a variety of social … structures which sociolinguists have typically not addressed", a position that acknowledges that language policy cannot occur independent of the social elements and factors within the target community. Ó Riagáin's (ibid) argument highlighted the advantage that historians have, when assessing the impact of historical events on language practise, that the outcomes, i.e.: language change, have already occurred and such identifying long term change is easier. However, current studies and researchers do not have this benefit of hindsight and as such need to be versed in many different disciplines, such as political science, economic ideology; with the additional sciences, researchers and social commentators must try to identify and

classify which social shifts are significant in the long-term as well as the shorter term, a challenging task that researchers in language policy are faced with.

As such Wright (2016:3) defines that sociolinguistics within language policy and planning needs to be inherently "interdisciplinary," which does present issues and risks. For instance, it is important to remember that language factors and language change are not exclusive of social change in the wider area. For instance, language practises can be affected by influxes of people into a smaller area, thereby increasing population density (Skeldon, 2009). Political and social change between communities within a singular country or region can be reflected in language use and language attitudes, this is a common place phenomenon globally (Sankoff, 2016).

Language attitudes can range from mild differences resulting in subtle biases that do not impact on a person's life (Schoel, 2012; Bassiouney, 2014), to nationalistic and extreme level differences that result in schisms between populations (Jones & Askew, 2016). Examples of the lesser extent can be seen in the North-South divide of England, whereby the southern varieties and phonological features thereof, are considered more desirable for economic and social prosperity (Levon et al., 2021); a result of both language attitudes following a minor language policy in media, which favoured the southern-linked accents, as well as the socio-economic power of the southern regions; as the government was (and still is) based in the south (in London) and the majority of the wealth was located In the southern regions, such as Kent and Buckinghamshire (Smith, 2013) . The attitudes found in England highlight the connection between the economic status and social mobility of a language community and the language practices in use across said community; in this case the

prosperity of the south encouraged the northern communities to adopt the language practises of the south; to improve their social status and economic prosperity. The case of England is lesser, in terms of attitude differences because the change was peaceful and occurred gradually (Crystal, 2003). There are elements that occurred within the social change that were not connected to the specific linguistic changes or uses, such as companies locating to the southern regions from the northern regions (Ibid). The movement of companies in this instance was related to the trade access of the southern region being greater than the north; however, the results of the movement were reflected in the language status quo and attitudes (Peled & Bonotti, 2019). Whilst the decision as to where companies operate from is not linguistically based, or even primarily involved with the language practises of England directly (Crystal, 2003), the results are connected directly to the linguistic attitudes as there was language change that occurred as a result of the population migrating south in order to maintain employment (Levon et al., 2021). This example, which is still ongoing, is used to expose and clarify that language change is directly connected to social change within a peaceful and co-operative manner.

An example of a stronger, almost schism-like, difference between a community based on linguistics attitudes and language use can be found in the Bedouin tribes across the desert regions in the Arabian Peninsula and Northern African regions, who have historically treated non-Bedouin speakers with contempt (Bassouniney, 2009). In the Golden era and onwards (est 600CE-1400CE) (Chejne, 1968; Cadora, 1989), the Bedouin held strong status in the Arabophone and Islamic nation states, as the nomadic tribes were considered the holders of tradition both historically, but

also linguistically (Versteegh, 2014). It was standard in the Golden era for the Bedouin to be considered the correct speakers of pure Arabic (by virtue of the prophet Muhammed, being a member of a Bedouin tribe) (Versteegh, 2013), and deviation from the Bedouin's speech community was to distance oneself from not only the tribe, but also from the status quo (Chejne, 1968). The Bedouin communities were historically conservative, favouring the retention of practises and methods from previous generations, and so to deviate linguistically was seen as an act of deviating from the conservative social state and community (Versteegh, 2013).

To maintain a stable linguistic identity, and to avoid any internal linguistic change, the Bedouin were both isolated, often avoiding all but essential contact with other communities, but also during events of contact, they were discriminatory to outsiders (Mosco & Atzaba-Poria, 2016). When imams and scholars travelled to the Bedouin tribes to learn and research, it is widely accepted that there were severe laws and regulations regarding the following of Bedouin speech (Cadora, 1989; Versteegh, 2014) with punishments for any deviation from the appropriate way to pronounce Arabic around the Bedouins, particularly the leaders. This highlights how a language practice, and policy, can be used to directly divide two or more communities, in this case the nomadic populations from the urbanised ones. Even in the modern day, the Bedouin tribes are still isolated by choice and resist any change in linguistic convention, as a method to avoid any major social change in the tribes themselves (Versteegh, 2014).

It is worth noting that the historical impression of nomadic tribes as negative and combative is not fully founded, as there are many instances across history and in the

modern era where the Bedouin have accepted outsiders and allowed for variation in speech (Versteegh, 2014). However, these interactions are often the result of compromise and considerations on both sides (see Tubi & Feitelson (2016) and Mosco & Atzaba-Poria (2015) for examples of outsiders working with the nomadic Bedouin communities). Furthermore, there is variation across the Bedouin tribes in the extent to which the negative language attitudes are retained and utilised; for instance, the tribes in the Arabian Peninsula are considered more accessible to researchers than those of the Maghreb regions in Africa (Bassouniney, 2009).

Furthermore, it is noted that whilst multiple disciplines are a needed element of holistic sociolinguistic research, there is a greater risk of discipline specific biases interfering with the research (Wright, 2014), such as an Islamic-focused historian favouring texts from Islamic scholars in regards to language use in the Caliphates; under the guise of the legal system of government and state were Islamic and thus the Islamic scholar would understand, and therefore report on, the status quo more so than non-Muslims or secular historians (Versteegh, 1993). This notion highlights how the political or religious history can be biased towards supporting a specific approach, which was produced to fit an agenda (Versteegh, 2014). It is not always the case that Islamic scholars would report on historical language use more accurately than secular scholars, based solely on the religious beliefs of the scholar. Yet, when analysing history, with that idea in mind, the results will be skewed in favour of promoting Islamic-based studies and approaches towards the Caliphates as opposed to secular. The issue of discipline bias is particularly brought forward when sociolinguistic research is cross-disciplinary with political science and other

ideological and philosophical sciences and subjects. An interdisciplinary approach is supported by Wright (2016:3) yet she warns of the riskiness of going "beyond one's own training" but also of the riskiness of "not do[ing] so;" a moot point that illustrates the complex nature of language policy and language planning, a sentiment that was presented by earlier sociolinguists, such as Halliday (2001:177) who stated, "language planning is a highly complex set of activities". Ergo when considering language policy, with Wright's tactile warning of overstepping one's own boundaries and knowledge, it is vital to provide a clear scope and rationale for the elements or regions investigated.

1.2 Scope of research

The following section will outline the scope of research for this thesis; the nature and elements involved with a specific set of environments, features and philosophies. The focus will be on analysing language policy and language use in situations and circumstances where the status quo is currently showing major weaknesses that must be addressed to improve the lives and experiences of those in need. This approach will follow the ethical consideration from Kabranian-Melkonian (2015), which identified that those populations which are at greatest risk from the negative effects of either mismanagement or instability, should potentially be the population from which sociological factors are assessed on. It is important to emphasise that this approach uses a risk-based model, in that the higher the risk, the greater the attention should be given, a sentiment repeated across emergency or disaster-based research (Sandman, 2012).

In practise, for language policy research, a risk-based model can support investigating minority languages in areas where there is opposition to the language community, or to the use of the minority language itself, such as, when a minority language is either: banned, deemed socially unacceptable (and sometimes instigating violence) or is a negative factor for socio-economic prosperity, such as being fired for speaking the minority language in the workplace. Another situation that is prioritised in risk-based language policy models can be found in emergency situations, such as natural disasters or civil disorder; whereby the focus can be placed on investigating and assessing whether the at-risk population would be able to access information in a language and format that they can understand (further discussion of disaster language policies can be found in section 2.3.2.2).

The latter shall be the approach of this thesis, based on a recommendation from language-based Non-Governmental Organization (NGO) Translators Without Borders (TWB), who identified that: "...further systematic research is required to inform communication strategies. This should measure actual comprehension and its impact on choices. It should confirm the most effective languages, formats, and channels for listening to and informing refugees and migrants from various language groups." (TWB, 2017a: 15).

Whilst the scope of this research is overarchingly motivated by this statement; further, the specific effects, events and situations discussed and presented in TWB (2017b) will motivate sub-questions and specific areas of investigation. This piece will aim to address, explore, and investigate three of the topics addressed in the overarching statement above. The first section explored will be that of the question

on the most effective languages to use when informing populations in situations of emergency, displacement, or disaster. The second section of comprehension and intelligibility shall be analysed, again, with focus placed on the processing of language in disaster-stricken and displaced populations. The third (and final) section will be that of the status of language attitudes and the impact that said attitudes can have on individual responses and chance of compliance with evacuation orders. Therefore, the overall thesis contributes to the "systematic research" mentioned in the TWB quote above, both from the perspective of measuring actual comprehension (second section) and impact on choices (first and third section).

1.3 Translators Without Borders Report – State of Affairs in Southern European Entry Points

The regions and societies of North Africa and the Middle East have been experiencing widespread instability in this millennium; both politically and socially (Zuber, 2018). In 2010, the Arab Spring began, with social unrest in North-Western African countries, such as Tunisia; the phenomenon of social unrest spread east affecting neighbouring countries, for example Libya, Egypt, and Syria (Dankert, 2017). In some nations, the social unrest resulted in conflicts between established governments and rebellious factions; notable examples include the Libyan Civil War [2011] and the Syrian Civil War [2011-]. A resulting effect of these conflicts was the mass-movement of people, with the populations of the waring country fleeing the conflict zones, in hope of safer environments and livelihoods (Dankert, 2017).

Popular destinations for refugees are the coasts of southern Europe, as the Mediterranean Sea connects Africa, the Middle East, and Europe (TWB, 2017a). This sea-route is used both legally and illegally (Dankert, 2017), to transport refugees from Northern African nations, like Tunisia, to either Mediterranean Islands, such as Malta and Crete, or to coastline countries, such as Greece and Italy (TWB, 2017a). There have been over a million migrants that have landed on the shores of Southern European Nations (Koroutchev, 2017); a phenomenon that has been labelled, by the United Nations Commission on Human Rights (Berry et al., 2016), as a migrant crisis as the communities, and their facilities, that are experiencing the influx are ill-prepared for a massive increase in population (Ibid). The problems are two-pronged; firstly, the regions were caught off-guard, due to the unexpectedness of the refugee crisis, especially the scale of the crisis, and secondly, the logistic operations to prepare and cater for refugees are resource heavy (considering the costs involved building enough shelter for over one million people), and there is a lack of political will to redistribute state resources for an over-haul project (Balkan & Tumen, 2016).

In practice, the refugee entry points are supported and operated by a combination of government agencies, such as border forces, and non-governmental organisations, predominantly charities (Aguistin & Jorgensen, 2018). All parties involved provide a form of aid to the refugee population (Fierros et al., 2017), whether that be aiding asylum applications by providing official documents to prove age or name, or to provide basic shelter, food, or medication (ibid). However, it is important to note that these organisations and agencies often operate independent of one-another, and as such the rules for aid delivery vary just as greatly as the deliverer of the aid

(Broussard, et al., 2019). A defence for the ad-hoc approach to the situation is borne out of the sheer number of issues being faced in the refugee camps; Agencies and organisations must prioritise on which issues they will fix and when (Maestri & Monforte, 2020: Gerver, 2021b), given that fixing all issues is too high a bar for the resource limited agencies.

Another issue facing all agencies and NGO's is that of the language diversity within the refugee population, given that not all refugees originate from the same point, and thus, they do not share the same language. Communication is a vital element for co-operation, whether that be between the NGO's themselves, NGOs and the refugees or even within the refugee population themselves (Tanner & Obrecht, 2015). Given the issues, it is not surprising that the agencies are currently using an ad-hoc interpreter network. TWB (2017b) reported that the international organisations, such as Oxfam, have acknowledged that there are communication difficulties with aid delivery, ranging from a lack of interpreters or difficulties understanding speakers of different language varieties. This is important as there was also an admission that the aid organisations were not addressing this issue of communication barriers. While this is justifiable, given that the agencies are prioritising the delivery of the basic necessities of life, the issue of communication barriers is an important gap, which this thesis aims to address.

The NGO Translators Without Borders is focused on providing disaster areas with language support to reduce the strain on the interpreter network, whilst also improving the quality of the interpretive services (TWB, 2017d) and eliminating language barriers within emergency situations (TWB, 2017b), such as refugee crisis'

(TWB, 2017a; 2017c) and disasters (2017d; 2017e). Whilst providing support, TWB operators are also recording and analysing the status quo of the linguistic landscape, diversity and issues found in the aid camps (TWB, 2017c). These reports have highlighted the impact the linguistic difficulties have on the overall process of aid management, as well as on those involved in the situations; including the aid workers, military officers and refugees (Tanner & Obrecht 2015; TWB 2017b, 2017c; 2017d).

In a 2017 report TWB (2017a) assessed the language situation of the southern European migrant crisis; investigators interviewed refugees and aid workers at two sites: Sicily, Italy and Chios, Greece. Both investigation sites were islands which had received refugees via the sea; during 2017 nearly one hundred thousand migrants arrived in Italy, 95% of which travelled from the Libyan coast (Italy Sea Arrivals Dashboard UN, 2018). Between May & June 2017, TWB (2017a) detailed a rapid assessment of the language support at the two entry points; with the aim of providing a snapshot of the situation to allow for research to be conducted to address the situation. The TWB report (2017a) contains verbatim accounts of the interviews with either refugees or aid workers, which can be further analysed to provide foundational depth and insight into the issues being faced in the crisis, since the report was designed to provide a springboard for future research. The TWB report (ibid) acknowledged that depth and analysis was not their aim. As such, the TWB (2017a) report identifies a research gap but does not provide suggestion for how to analyse or improve the situations faced in the field; this thesis is motivated from said gap identified.

With the research gap of communication issues in the field being identified, the following section will begin the work for analysing the TWB report, with the intention to identify further avenues of research and consideration, to address the overarching research gap of communication issues in the field of refugee and emergency management.

1.3.1 Reviewing TWB

The first issue of language use and planning in the field, that was identified from the TWB (2017a) was that of Language Support, which regards the physical and logistic support for languages and communication. Currently, the support level for refuges is highly variable, with some areas having greater support than others (TWB, 2017d) . For instance, if a South American refugee is in a more developed country such as the USA, there is greater access to translators (Gerver, 2021a). This contrasts to Southern Europe, where TWB (2017a) calculated that 1/13 refugees had access to an interpreter during the processing period of registration. This is a critical issue, as 12/13 of the refugee population are not being supported; thus, their integration into a new society or even into a temporary refugee camp is hindered. Language barriers are regularly recorded across many conflict zones, especially when there is a sudden influx of refugees, such as Chios, Greece, and Nigeria (TWB, 2017c). Thus, there is a slim chance that refugees can easily or efficiently register as a refugee (TWB, 2017a) to continue in the process of becoming a permanent residence of the host country. The registration stage of refugee management and processing is a critical stage (Kabranian-Melkonian, 2015), as inaccuracies or errors with documentation of processing paperwork can result in life-altering outcomes (ibid); ranging from

incorrect relocation placements (Fierros et al., 2017) to inaccurate medication information and therefore inaccurate treatment of pre-existing conditions, such as allergy information being missed (Clarke & Isphording, 2015).

To illustrate the difficulty faced by aid workers and operatives in processing migrants and refugees, the minimum information required to process a refugee is: Name, Age & Country of Origin (TWB, 2017). The three pieces of data are basic demographic information, which are considered essential to refugee processing (Kabranian-Melkonian, 2015). The complexity of the responses is low, as there is no need for the refugee to provide full sentences when replying; syntactically it is possible for the responses to be verb-less, such as 'Twenty' and 'Syria.' The vital element of the responses is the semantic information, rather than the grammatical information. However, it is unusual to speak verb-less sentences when discussing formal matters or in discourses (both formal and informal), such as the interview scenario for the refugee during processing (Sourander, 2003; TWB 2017a). Typically, there is a government agent and the refugee (and any accompanying family) (Segal & Mayadas, 2005); this situation is not typical for the refugee so due to the pressure, as passing the interview will result in the refugee/s getting help, and the unfamiliarity of the environment and the speakers, the overall discourse is held in a formal manner. A refugee when asked for their age (assuming that they understand the question), will provide full sentences rather than the basic 'broken' sentences; thus, the critical information needed is mixed with other linguistic pieces, such as grammatical function words and other semantic function words.

For instance, the following examples show how critical information can be blended with non-critical information:

A)   The interviewer hears the following response after asking; where are you from?

'I am Syrian but I need to go to France to be safe please let me through Greece'

The response contains three words that could be the answer to the question; Syrian, France and Greece; all three words are Proper Nouns, thus are semantic units within the sentence. The interviewer can identify the correct response to the sentence based on the grammatical information, or the functional affix; which is on the root Syria, as the '–n' indicates a relationship between the speaker and the semantic notion. Thus, the following options can be easily disregarded, as they do not contain any grammatically functional elements.

B)   The interviewer hears the following response after asking; where are you from?

'I am from Syria but I need to go to France to be safe please let me go through Greece'

There are three possible answers to the question that can be extracted: Syria, France, and Greece. Given that all three of the possible answers are Proper Nouns, they cannot be differentiated by grammar. Since there are no inflections that indicate a relationship between the speaker and the noun, for the interviewer to obtain the correct answer they must process the sentence in further depth, decoding the verbal elements, which change the options to the following relationships:

'to be' + 'Syria'

'to go' + 'France'

'to go' + 'Greece'

From these options it can be identified that 'Syria' is the country of origin, as it is accompanied by the copula infinitive, which indicated a relationship between the Subject (the refugee) and the Object (country of origin).

These two examples highlight the different level of processing and parsing required by the interviewer to obtain the correct response from the refugee. In example A, grammatical information is needed to decipher which of the presented nouns is the correct response, whereas in example B, the verbal information is needed to identify the country of origin. Thus, in comparison to broken sentences which only contain a noun, the full sentences take longer and require more processing, which allows for greater chance for errors to occur. During the initial stages of processing refugees, only basic information is required (Clarke & Isphording, 2015), such as Country of Origin or Name; which can be delivered using broken (or partially intelligible) sentences; since, linguistically, and semantically, the minimum amount of information required can be delivered using a single noun, or noun phrase for each response. However, as the refugee is processed through the system, the complexity of messaging also increases, i.e. at first just the country of origin is needed, when later assessment questions relating to how the refugee arrived, so method of transport, and the purpose for their relocation, or reason to put on the application for asylum status.

From the perspective of refugee management, the basic demographic pieces of information are the minimum as they enable the connecting or reconnecting of populations or families (Fierros, 2016). This is a primary goal within refugee management, as current research supports unifying individuals based on these their age and country of origin (Geiger & Pécoud, 2010), as these factors provide a social connector (Sourander, 2003) between the individuals. The name is less important when grouping larger numbers, such as processing boatloads at a time (Kabranian-Melkonian, 2015); however, it is vital to the re-uniting of families, such as parents with their children (Segal & Mayadas, 2005). All three pieces of information are used in tandem to identify and distinguish between the masses of refugees across a range of locations (Fierros, 2016). An error within these basic details can have a cumulative effect for the individual refugee whose file is incorrect (Kabranian-Melkonian, 2015). For instance, if all three details are errored, then the unification of the refugee with family members or close ones becomes difficult at best (Jastram & Newland, 2003). If there is an error in the recording of Country of Origin, then the long-term refugee management choices are affected, for instance, the refugee could be transferred to a region designated for a different culture or nationality (Fierros et al, 2017).

Additionally, with age errors, the most effected are teenagers, or pre-pubescent individuals, as the visual identifiers of adulthood are highly varied (ibid). For instance, if there is a teenager (aged 13-17), who is incorrectly identified as an adult, as they are similar in height to the average adult, this could have a knock-on effect in their relocation (Fierros et al, 2017), as they would not be given the correct support needed for children, who are generally more vulnerable than adults (Geiger &

Pécoud, 2010). This example of error can result in children being treated unfairly (Segal & Mayadas, 2005). The reverse is also possible, whereby adults who look visually younger are classified as children, and as such are given less freedom and independence (Hart, 2009). With both cases it is difficult for the individual to change their refugee classification, i.e., the details on the documents, as amendments typically require evidence, such a birth certificates (Jastram & Newland, 2003), which many refugees do not possess due to their rapid or sudden displacement from their Country of Origin. Another method to amend the documents is if there is a credible person who can vouch for the refugee, such as a family member; the chances of the individual finding their family members is heavily hindered by the document errors (Williams et al., 2016). Thus, the refugees who are incorrectly documented face additional challenges and difficulties than other refugees (Fierros et al, 2017).

There are noted communication issues that are occurring along the Southern coast of Europe (TWB, 2017a), which are identifiable, and recognisable from a linguistic perspective. However, the actions taken to address these issues have focused on alleviating the language barriers in the situation rather than understanding the causes (TWB, 2017a). The first step towards understanding the complexities of the current situation is to analyse and investigate the sum of the parts involved within the field. This thesis will aim to address part of the research gap, by investigating the difficulties in language planning for disaster events and emergencies, whilst also exploring the status quo of linguistic variation and language demarcation. Furthermore, this thesis will also consider and evaluate the status quo of social attitudes based on linguistic features. This thesis will aim to provide an overview of

primary issues identified from TWB (2017a); appropriate and understandable language selection within language policy to facilitate communication in a disaster event, from a linguistic stand-point. The next section will consider the state of art in linguistics and disaster research, to provide a foundation for the investigations and experimentation of the status quo issues.

Chapter 2: State of the Art

2.1 Identifying the multidisciplinary factors involved in planning disasters language policy

Given that this thesis aims to explore the linguistic issues within the communication of information in situations of disasters, it is vital to first limit the discussion from disaster response overall to the specific multi-disciplinary elements that interact with linguistics. To facilitate this approach, this literature will be separated into two sections: the periphery and the linguistic. The periphery sections will explore the contextual elements required to investigate disaster responses, to ensure that this thesis does not ignore the relevant issues that are unique to the disaster context. This clarification is important, as this thesis is applying linguistic theory to a contextual and multi-disciplinary issue, that of disaster responses, which involve psychological and humanitarian research to be considered to retain contextual awareness. The linguistic sections will explore the language-based issues that are relevant to the disaster response context, such as language use, or language variation.

When considering disaster communication, the first step required is identifying the overall approach taken in the development of the underlying language policy. There are two general approaches, based on the direction of change between the different levels of authority and community (McMenamin & van der Walt, 2018). Firstly, there are bottom-up policies, which are produced by a small community, and develop to influence language change to the wider population (Johnson, 2003). Example

bottom-up policies include the Welsh language revitalisation movement, which began with individual communities regulating that Welsh should be used instead of English in the domestic contexts, which led to social change supporting the use of Welsh beyond the home (Williams, 2000, 2014). The other approach is top-down language policy, when the decisions regarding language use, and which language to use, are decided by an authority who in turn enforce the language rules (Wright, 2016; Aboagye Da-Costa & Adade-Yeboah, 2019). A characteristic of top-down is that those at the bottom-end of the language policy do not have sway as to what is involved with the policy, whereas with bottom-up, individuals can impact the policy's development (Johnson, 2003). Given these distinctions, disaster language policies, when introduced and planned alongside disaster prevention policies, are consistently going to be top-down. Whereas, if there is no planning for a disaster, the language policies enacted will predominantly be ad-hoc, and therefore bottom-up, as the local communities would decide on language use, rather than a governing authority. This thesis is investigating language planning for disaster events; therefore, it is appropriate at this stage to limit the literature review to considering only top-down policies from this point.

2.2 Periphery factors for disaster contextualisation

2.2.1 Defining and categorising human disasters

This thesis will explore language use under adverse conditions caused by a disaster event. With that in mind, the first step is to define what a disaster is, as the definition is disputed. The debate regarding the term disaster itself has evolved since the early

1900s. Currently there is a debate relating to the theoretical elements that are involved in a disaster, or disasters in general (Perry, 2007). As such, there is no agreement on a specific working definition, and there is debate as to what should be included and excluded from being part of a disaster (Lindell, 2013). Perry (2007) criticised that the efforts for defining what constitutes a disaster 'brings out the pedantic in scholars' (pg.1), however, Perry (ibid) also does not advocate for the promotion of an arbitrarily simplified definition, rather they support finding a balance between detail, depth and usability of a single unified definition. The criticism is significant since it is agreed that disasters majorly impact human populations and result in harm and injury to those involved, so to fixate on the theoretical constructs disregards the real-world situations and contexts of disasters. This thesis will present disaster research and applied standards, focusing on fieldwork-based situations and previous disasters, rather than focusing on the theoretical constructs as to what could be potentially defined as a disaster.

With that theme in mind, the definition of disaster will be required for working use in this thesis. The following working definition presented is a combination of definitions and characteristics from previous disaster research scholars, including Perry's (2017) inclusion of social disruption as a key feature of disasters, Lindell's (2013) considerations of situations with sudden collective onset of stress, and Carter's (2008) overview of disaster management requirements.

Working definition of a disaster for this thesis:

An impact event which causes damage, harm, or an increased risk-to-life to the population, with direct consequences of increasing pressure and stress on the civic

infrastructure, social framework and/or the ecology, all culminating in a reduction of

overall human stability of the impacted population.

This definition is for disasters overall, as a universal category of event. There are

three overarching groups of disaster, based off the style of the impact event. These

groups are Human, Natural or Hybrid (Quarantelli, 1991; Perry, 2007). Natural

disasters are those which are caused by naturally occurring phenomenon, exclusive

of human activity; examples include: earthquakes, because of tectonic movements

(Nakanishi & Matsuo, 2014). Human disasters are when an impact event is a result

of direct human-based actions, exclusive of naturally occurring phenomena,

examples include mass violence (Bolin, 2007). Hybrid disasters are when the impact

event is a result of both human action and natural phenomena (Boyarsky &

Shneiderman, 2002). An example is the disaster event of Fukushima 2011 (Okada et

al., 2021), when the Tōhoku earthquake resulted in a joint disaster event as there

was the tsunami alert alongside a nuclear meltdown alert. The overall situation of

Fukushima is hybrid due to human and natural factors creating risk events to the

population. This thesis will consider human-based disasters only, this is justified by

the context of the motivational situation for the project: the Southern Europe

Migration Crisis (TWB, 2017a; 2017c), which is a consequence of human conflict,

which is thus a human disaster situation.

To investigate human disasters and allow consideration of the elements and

differences between the sub-categories of human disaster events, this thesis will

present four measurements to distinguish between different disaster events and

provide an objective system to allow for like-for-like comparisons further in the

project. These measurements have been motivated as being part of either the impact event itself, or the repercussions of the event. Furthermore, the four measurements account for a separate element of a disaster, which in combination can provide a summary of the core features of each event. The measurements are, as follows, Cause, Environment, Geographical Size, and Impact (Quarantelli, 1991: Shaluf, 2007: Vij, 2022) . The following section will outline and describe the levels of each measurement, providing real-world examples for reference.

2.2.1.1 Measure 1: Cause

There can be multiple causal factors or events which transpire either simultaneously, or in-continuity, which develop into a disaster (Etkin, 2014). This thesis will be limited to examining either factors that occurred in union, or in the case of continuousness, the latest three previous events that resulted prior to disaster. Given that disaster episiotomy has been criticised for being never-ending with unnecessary detail, depth, or considerations (Perry,2007), it is important to limit the additional factors that can be included in any disaster research. This can ensure that the focus point of research remains on the specific disaster event, and is not distracted, or muted by the events that the disaster developed from.

As this thesis aims to investigate the effects and management of language use and communication following a disaster, then the limitations must also follow suit. Therefore, three primary causes that will be considered in this thesis, based on Kreps & Bosworth (2007) and Carter (2008): A) Escalation B) Intentional attacks and C) Accidental. The following section will outline each element.

A)   A disaster which begins as a minor or lesser emergency, which then escalates upwards in terms of severity or risk-of-harm. An example situation which can escalate into a disaster are wide-spread protests, which can escalate into violent disorder. Examples of protests that escalated into violent disruption include Benghazi in 2011, where protesters were fired upon by the security forces and as a result a civil war began (Aghayev, 2013). In the Benghazi situation, the limit of three previous events allows for the political situation of Libya from around 2008 to be included, from which the root cause can be identified as political instability and a civil rebellion brewing due to widespread economic disparity and a political system that favoured discriminatorily between the people (ibid).

B)   There can be premeditated disasters which are when the impact event is designed to cause maximum harm of disruption to a population or area. Terror attacks fall into this classification, as they are designed to cause damage to a populated area, injure the local population, and encourage a sense of panic or terror in the wider community. Examples of high impact terror attacks include the *2001 9/11 Twin Towers* and *the 2011 Utøya* attacks (Dankert, 2017); which were unexpected and caused wide-spread damage and distress.

C)   Accidents can happen due to infrastructure failing, human error, miscommunications, or situational-based miscalculations. Examples of accidental disasters include chemical explosions or industrial leaks, such as the 1984 Bohpal disaster (Pietersen, 2013) and the 2000 Enschede firework disaster (van der Velden et al., 2012).

This thesis will focus on Escalation disasters (A). As the Arab Spring protests, and subsequent civil unrest and conflict, was the cause for the mass migration into southern Europe (Zeber, 2018). As such, this thesis will use this measure to distinguish between literature and case studies on disaster events.

2.2.1.2 Measure 2: Environment of event

Two factors that need to be considered when preparing language policy for disaster response are land infrastructure and land use. Language planners need to account for the different locations that the population can be in, and the development of said locations. For instance, humanitarian relief relies on roads, thus the development of transport infrastructure should be considered by disaster planners (Pettit et al., 2015). Within infrastructure there are two groups, Urban and Rural, whereas with land use, there are three sub-categories: Residential, Commercial and Industrial. The following section will briefly outline the different need for each category in order of appearance.

In general, an urban area is characterised as a region where the population is concentrated into small, built-up centres, such as cities and towns (Britton, 2007). The additional characteristics are high-population density, widespread evidence of human development and civic infrastructure (such as roads, sewage and electricity). In contrast, rural areas are characterised as being remote regions which are sparsely populated across a large geographic space, therefore having low population density, as well as limited civic infrastructure (such as single roads between settlements) (Nakanishi et al, 2014).

This distinction is relevant for language policy in a disaster, since urban dwellers are modernised, using technology to communicate in the day-to-day life, whereas rural populations favour traditional technology such as radio or television (Demuth, 2002). When considering communication in response to a disaster event, the infrastructure does impact the effectiveness of the communication strategy (Nettle, 1996). To illustrate this, consider a mobile SMS alert system to warn the population of an incoming disaster, a strategy reliant on mobile phone usage. The density of phone ownership differs between urban and rural areas, with the rural typically being lower (Demuth, 2002). Therefore, the SMS alert system would have a lower chance of success dependent on location. Human development and technology use are connected (Mayhorn & McLaughlin, 2014), and disaster language policies must not assume that all nations have equal access to technology, particularly in the developing world, where technology outside of urban cities is still in its early stages (Vij, 2022).

With land use, residential areas are those which contain human dwellings as the primary function of the land (Quanantelli, 1991; Carter, 2008). Industrial use is when resource production and manufacturing is the main function (ibid), and commercial use are areas of trade and commerce as the primary function (ibid). The disaster response for a region is influenced by the land use.

With residential areas the primary aim of language use is to validate the severity of the disaster, in terms of risk, to convince the population to evacuate (Kuligowski & Gwynne, 2010). Furthermore, language is used to instruct and co-ordinate evacuations, particularly at night-times, when ordering residents to leave without

taking unnecessary items with them and reassuring the population that the safest course of action is to evacuate (Inoue, 2012).

In industrial areas, the primary aim for language use, is locating at-risk populations for evacuation (Doğan, 2016) given that workers are dispersed across a low-density local area, as the infrastructure is built for large-scale use. Also, the location of workers can vary, the staff operate across the wider industrial campus. Whereas, in a residential area, the population can be estimated based on the number of houses, however, the same method cannot be applied to industrial areas, as the building size or density provides little correlation to the number of workers inside (Quarantelli, 1991). Therefore, communication needs to focus on encouraging workers to direct disaster response operatives to the sub-sections where there are people.

With commercial areas, the foremost task for language use is to alert the population without causing a sense of panic that can result in civil disorder and erratic en-mass fleeing (Calhoun, 2008). The commercial areas are designed for traffic flow, both of vehicular and pedestrian, and these systems can become overwhelmed by evacuations, particularly during a panic when people run (Maheshwari & Rajan, 2016). A common mechanism is using a public announcement system (Donner et al., 2007; Arai 2013b), which relays a single message across a wide region, such as inside shopping centres; however, the effectiveness of these systems is dependent on the delivery of messaging (a factor explored in depth later in *2.3.1.7 Guidelines for emergency communications*).

This thesis will focus on the residential and commercial environments, as the environments reflected in TWB (2017a) are temporary aid camps, which are

comprised on residential tents and within the campus's there are local businesses, set up by either the refugees or the local populations.

2.2.1.3 Measure 3: Size of event

Disasters can be categorised according to the scope and size of the disaster both in terms of the initial impact event and the area experiencing the resulting knock-on effects (Perry, 2007). Within research and disasters there is little consistency as to defining the different size levels of both parts of the effects (ibid). It is generally agreed this scale is a continuum, with the highest effect size being that of the total population of the planet, whereas the lower level of the continuum would be localised and impacting a small group of individuals (Carter, 2008)

To categorise disasters by size, this thesis will use the politically designated regions as the measure, such as legal jurisdictions. There will be four groups: Localised, Regional, National and Global (Carter, 2008). Therefore, language policies will be sorted into the four categories based on the following geo-political criteria:

1) Localised – with community level jurisdiction. Examples include geographically isolated communities such as village councils in Sri Lanka or the council of elders in the Saharan Bedouin tribes.

2) Regional – with regional jurisdiction covering multiple communities. Examples include autonomous municipal areas, such as Catalonia in Spain, and semi-autonomous unitary authorities, such as the Greater London area in the UK.

3) National – with governmental jurisdiction of a whole sovereign territory. Examples include any nation recognised by the United Nations (UN).

4)   Global – across multiple national jurisdictions, not necessarily with direct authority over the nations. Examples include the European Union (EU) and the United Nations.

This thesis will focus on the events which are regional or greater. A decision motivated by the TWB (2017a), as the Southern European Crisis is a result of regional, national & global human conflicts. As the focus is on investigating the overarching issues facing the humanitarian and refugees sectors, limiting to the localised level would reduce the scope below requirements.

2.2.1.4 Measure 4: Impact

Past experiences and examples are used when developing policies for disasters, allowing for past-mistakes to be avoided, and successful elements to be maintained. This process is reliant on there being A) similar disasters in the past and B) a system for quantifying and categorising the similarity of disasters. Given that this thesis is assessing language use in disasters comparable to the Southern European Crisis, similar cases of past disasters need to be identified. In this case the criteria will be firstly, disasters which result in population displacement from conflict regions, and secondly, resulted in casualties. These two criteria can be measured by the following:

Displacement – Volume of relocated individuals, as a raw value and as a percentage of emigration from the conflict zone. Within this, there are three subgroups: Forced, when the relocated are not allowed to remain in the at-risk zone (Jacobsen & Landau, 2003); Coerced, when the population is encouraged through threats to leave the danger zone, this can be using direct negative threats, such as attacks, or

through sociological methods, such as threat of isolation (Hart, 2009). The last sub-group is voluntary, when the population leaves without being ordered to, or instructed to (Balkan & Tumen, 2016). This latter group also includes people who leave an area when they feel at risk or are struggling to cope in the at-risk area.

Casualties – Volume of fatalities attributed to the disaster, both during and post-event. As a caveat, calculating and comparing with this method should only be deployed with sensitivity (Kabranian-Melkonian, 2015).

This thesis will focus on the population displacement, as the situation in TWB (2017a) depicts the processing of refugees; those who have been displaced. Part of the issue facing the aid infrastructure is the volume of refugees that arrived, and therefore, comparison of refugee numbers is satisfactory.

2.2.1.5 Refugees & Demographics

As displacement is a core element for the contextualisation of the Southern European Refugee Crisis, the impact of refugees on the demographics of a region is warranted. As policies made using current demographics of a nation will need adapting with mass scale displacement (both with immigration and emigration). Therefore, this section will outline the three primary ways the demographics change with human disasters, these are: mortality rate, mandatory relocation, and exodus. The following section will outline each measure.

With human conflict, the areas involved become warzones, which increases risk-of-harm or risk-to-life. In the warzones, the mortality rate changes upwardly in comparison to the pre-war figures (Besançon, 2005). Whilst this is a morbid

measure, it is a reality of the disaster event of human conflict. For language policies, what is particularly important is the mortality rate in the minority language populations, as a reduction in speakers in smaller communities is more acutely felt in comparison to majority languages (Wright, 2014).

If the disaster management can support relocation of a population, through mandatory orders and governmental support, the demographics in the disaster region will change (Carter, 2008). The population in the disaster zone reduces, and a safer region will experience an increase. When this relocation is planned, the safe zone will receive provisions to cater for the refugees (Fierros et al, 2017), which include housing, food and medication as well as providing language support for the relocated population: the aspect which language policy is responsible for (Gardner-Chloros et al., 2016).

Mass unplanned exodus is population movement from a war zone into another area, of which the safety level is perceived as higher (Holmes & De Piñeres, 2011). This is most common in conflict situations when the fighting becomes guerrilla warfare in urban environments (Gerver, 2021b). The population evacuates the areas through whatever method possible to survive, as such the movement is erratic and dynamic throughout the process. As such, the relocation areas have little idea as to which demographic groups have been received, hindering the deployment of appropriate resources to support the influx of people. In these cases, communication between aid workers, emergency services, and refugees is vital to keep the tensions low and stable (Fierros et al., 2017), both within the refugee population, and between the refugees and the emergency workers.

The context of TWB (2017a) was in the processing of refugees in a mass exodus, and the report highlighted that communication failures can result in severe repercussions which are difficult to resolve. In processing refugees, there are five identified discourses where communication should be supported, to avoid misunderstandings. They are:

1) Directions, to control the flow of the refugees to the appropriate area or zone (CDC, 2014), to prevent chaos or confusion across the group (Gerver, 2021a).

2) Processing refugees upon arrival, the collection of basic demographic details, such as name, country of origin, method of arrival, nationality, reason for fleeing previous region or area (Fierros et al, 2017), to quantify humanitarian aid required (Pettit et al., 2015).

3) Evacuation orders, to provide an understanding of the severity of risk that can occur if a refugee is staying in an unsafe area or unstable environment (Britton, 2007), to counter dis-information as to the severity of the situation (Keselman et al., 2005).

4) Providing medical care, this includes the dispensing of medication, as knowledge of allergies, or ailments is needed prior to the delivery or application of treatment (Tanner & Obrecht, 2015), to prevent medical incidents and negative biological responses to medications prescribed (Williams et al., 2016).

5) Reassurances of safety, to calm fears and tension (Rajan, 2019), to support the healing processes for the trauma experienced by the refugees and

highlights the safe havens that aid camp are designed to be (Broussard et al., 2019).

Within this sub-chapter, the elements of disaster management that language planners need to consider have been outlined. This review was vital in providing the foundational and contextual backdrop for further investigation into the status quo of language policy and communication in a disaster. An aspect from TWB (2017c), that of language as a human right, was not considered as human rights are not an aspect for classifying and quantifying disasters. However, accounting for the language rights is still important (TWB, 2017b; 2017c), therefore the next sub-chapter will focus on language rights, and the connection to language planning for disasters.

2.2.2 Human rights and fundamental freedoms

Language rights are known as either fundamental freedoms or basic human rights (De Varennes, 2021). Whilst these terms may be different, the distance in meaning is little. A fundamental freedom is an ability or status of an individual that cannot be removed without violating an agreed code of human conduct (ibid). The term basic human right, likewise, when denied implies that an individual or group is being treated as less than a human being, by other human beings (Skutnabb-Kangas, 2006). Both terms denote the minimum level of protection that a person can enjoy uninhibited.

Human rights are enshrined in international and national based charters or statutes of legislation (Ammon, 2006). In the context of international based human rights the benchmarks are the United Nations charter and the Council of Europe charter

(Broussard et al, 2019). These two organisations are considered the guidelines for communication and legal minimum human rights, this is because the United Nations charter has been ratified by all nations in the organisation (Bourantonis, 2004). The Council of Europe charter is also a guideline because it specifically focuses on human rights and the importance of specific rights to improve humanity overall (Sokolovska, 2017). However, it is important to note that not all countries subscribe to the Council of Europe and as such they are not legally, or morally, bound, to follow the guidelines and principles that the council promotes (Ammon, 2006).

In contrast human rights can also be enshrined in national constitutions or legislations (Skutnabb-Kangas, 2006). When considering national based human rights, the overarching theme is variation. There are multiple avenues of human rights and laws which are contrastive between regions and nations, a clear example falls on the death penalty for criminal prosecutions, with some nations still holding the practice in their legal systems; whereas other nations have banned the practice claiming it is against their human rights for their citizens (Barnett, 2003). Whilst this is not directly related to language, it is important to recognize that the legal system for human rights and access to human rights is highly varied across the globe, thus assuming that all human rights of one country are therefore the same human rights in another is problematic.

The distinction between national and international human rights guidelines, and whether a set of human rights guidelines are applicable is vital when considering international disasters, because whilst an event can occur across two nations it is important to address the situation both as an international event, as there are

multiple nations involved, but also as an event in both nations individually. It is important that the national level rules and regulations be followed, and the guidelines of the international community should not superimpose moral codes upon the individuals of a nation, particularly if those rules do not coincide (Broussard et al, 2019).

The origin for modern day human rights is rooted in the United Nation's charter of 1945, and subsequent charters by the international organisation since (Weiss, 2010). There are multiple international co-operations that have acknowledged the UN charter as a guideline for the construction of their own charters, for instance, the African Union in the opening of their charter, ratified in 1963, which was eighteen years post ratification of the UN charter, highlights the connection between the African Union charter and the UN charter (Organization of African Unity, 1963).

With regard to language, and the human rights of language use, the United Nations charter has two articles that directly refer to language in the context as a human right or fundamental freedom (UN, 1945). Article 55 highlights that there must be mutual respect and observance without distinguishing by language, followed up with article 76c where human rights and fundamental freedoms should be encouraged, to recognise the inter-dependence of people. As previously established, language rights, as per the UN, are fundamental freedoms alongside race, sex and religion. These are the two specific references to language use in the UN Charter, and what they indicate is that language spoken should be a characteristic considered without discrimination; meaning that a person should not be given greater or lesser treatment by virtue of the language they speak or can otherwise communicate with.

The practicalities of applying a human right based guideline, where no individual can be discriminated or treated with alternative provision by virtue of their language, are resource demanding (Skutnabb-Kangas, 2009). It is an accepted compromise between individual fundamental freedoms and the functioning of the internationalisation, that there is a disconnect between the desired and the actual, in regard to how languages' are treated. Namely that some languages are prioritised above others, in contrast to the sentiment that all languages should be treated equally. Altruistically, all languages should be considered equally, but the practicalities result in there being an imbalance between majority and minority languages. This has resulted in there being a wealth of teaching resources for the six official language is of the United Nations, but less so for the other languages of the world (Johnson, 2013). Of course, it is important to remember as well that there are over 7000 languages in the world (Ethnologue, 2022), so to expect any organisation to be able to provide universal communication without any distinction is a goal which currently is not achievable (Dimmendaal, 2008).

When choosing which language to use and assessing language choice, there are supplementary logistical issues, namely: language resources, stability; and accessibility. There is always a limited number of resources that can be given in socio-linguistic situations (Androutsopoulos, 2014), so language planning must remain strategic. Logistical difficulties can manifest from providing support to minority language communities, minimal teaching resources and/or lack of interpretative and translatory infrastructure within a target region (Pettit et al, 2015). Managing and addressing deficiencies in language resources prior to policy implementation is

difficult to achieve, since 'language is both a personal... and a physical resource' (Ruiz, 2016 pp. 38). Speaker numbers are a typical baseline for logistic planning for multi-lingual communication across a target community, a baseline that is fluidic in nature.

Language resources range from formal teaching materials, such as textbooks, to informal materials, such as interpreters (O'Brien et al., 2018). Historically, sociolinguistics has focused on measuring resources through the education sector, such as textbooks for teaching a foreign language; partly due to the education being a primary tool in language revival and maintenance (Cabrera, 2014); and due to the easiness in which numbers can be acquired by researchers and government officials alike (ibid). Counting physical media resources is easier than recording human-based resources, such as interpreters and translators (Taibi, 2016), again attributive from the fluidity of speaker numbers.

When considering the resources available, it is important to evaluate the type of policy that is being assessed, and what is the driving force behind the policy's creation, such as political or sociological. According to Wright (2016), language policy begins with a core set of principles or themes, such as one nation unity or protecting historic groups; these principles are the foundation for any development of policy and as such are innately integrated into the assessment, evaluation & delivery options for policy in practice and in planning. When policy changes are advised, or alternative policies for replacement are suggested, the defect authorities on the matter, typically governments, view these changes as either "an opportunity rather than a threat" (ibid, pp. 11) or the opposite.

This argument builds upon Ruiz's (2010) orientations within language policy as a distinctive linguistic paradigm. These orientations relate to the position of entry into language policy for all discussion, analysis, or revision of status quo policies. Hult & Hornberger (2016) summarise Ruiz's three primary orientations, entitled as 'Language as 1) Problem, 2) Right & 3) Resource' (pp.33). Language as a 'problem' includes assessing language policy for a monolingual scope, predispositions that bilingualism reduces academic achievement and produces cognitive difficulties & minority languages as a threat to the status of the dominant language. Language as a 'right' relates to the individual and civil based legal right within a country, such as the ability to use a language in education, voting, or media; also, the right to speak any language is contained within this orientation category. Language as a 'resource' encapsulates language having value and purpose for official and unofficial activities, such as the impact of language policy in democracy and public relations. Language as a 'resource' also accounts for language resources, language based-social change or associations, such as negative stereotyping (Schiffman, 1998), resulting in a schism in a society developing into ethnic divisions within the same society (Pattanayak, 1988).

It is important to note, that the discussion of language resources in this text is aligned with Hult & Hornberger's (2016) orientation of 'language as a resource' (see pp.33 & 38-42). In contrast, Wright's (2016) proposed core themes of language planning are aligned with Ruiz (2010) and Hult & Hornberger's (2016) orientation of 'language as a problem' (pp.33). This is not to discredit Wright's (2016) interpretation of language

planning and accounting for resource management; it is to highlight different

theoretical origins within this discussion.

In situations of disaster, human rights are of vital importance but are suitably harder

to maintain (Agustín & Jørgensen, 2018). This is present particularly in sudden

disasters, where the impact event was unexpected, and the emergency services are

overwhelmed (Maestri & Monforte, 2020). In the event of a major disaster event, the

aims of the aid workers or emergency services is to relocate the population that is at

risk, provide basic medical support to any injured and to provide safety for the

population that has been relocated (Van der Borgh, 2009). In this situation, making

sure that human rights in relation to prohibiting any form of discrimination is difficult to

achieve, not because the aim of the emergency services is to avoid human rights,

rather that the situation is highly dangerous, and the saving of human life is a greater

priority than the rights of the individual.

The guidelines can be assessed depending on the perspective covered, as disasters

are multifaceted (Barnett, 2003). With language rights, the 4-A language framework

analysis tool is a way to show the mechanism to improve the language support

offered and the services thereof (O'Brien et al., 2018). The guidance's are not legally

binding in a national sense, they are binding for international law, which is far more

lenient to deviation and crimes (De Varennes, 2021).

A note on terminology, the use of the term language policy to describe the

overarching requirements of minimums and topics of language used in a disaster

response are not always labelled as policy per se, rather some refer to them as

frameworks (O'Brien et al., 2018). In emergency planning and management research

the focus is on producing logistics-led reports and as such the distinction between policy and framework is tempered and collapsed into a single term: framework (Schiffman, 2012). To address this, in this thesis, the term language policy will include the elements found in frameworks, such as, techniques to use in messaging, so long as it impacts the selection of a language overall or similar top-down planning aspects.

The 4-A Standards tool was designed to evaluate the right to education, and was created by a UN facilitator, however the approach was adapted for the evaluation of communication strategies and language frameworks by O'Brien et al, (2018). Therefore, this design allows for comparison of four key elements, allowing for the identification of weaknesses of a framework, as well as highlighting the areas requiring significant, or targeted, improvement. The full standards are accessibility, availability, adaptability, and acceptability (UN CESCR, 1999) (see table 1 for summary of the standards). These four standards when used in combination are supposed to be able to provide an objective similarity between differing nations which have different language groups, different language population sizes, and different levels of development. The 4-A standards are a stable foundation for language framework assessment, as the standards can be adapted and expanded to fit the purpose or aims of a particular disciplines line-of-reasoning when investigating communication in disasters and emergencies. Furthermore, the 4-A standards tool allows for language rights to be consistently considered universally, thus supporting the development of language frameworks and policy globally to achieve the

overarching goal of open-availability in information in all languages globally, as a steppingstone.

Table 1: The 4-A Standards framework

Availability	Language rights need to be economically viable for the population, in a method that is accessible in an emergency.
Accessibility	Language support needs to be provided without disparity or discrimination between speakers, so that all involved can fully understand the information and communications.
Acceptability	Language and communication provided needs to account for the human rights of the service users, particularly by providing high quality resources and information.
Adaptability	Language frameworks and communication systems need to account for the cultural and linguistic diversity found in the affected areas. Also, frameworks need to be adaptable to ethno-demographic changes in the affected areas.

The 4 A-standards are adaptable and can be changed to fit specific areas or elements of disaster communication research. O'Brien et al (2018), used an adapted version of the 4-A standards to analyse the emergency language frameworks of five higher-income nations. The study focused on assessing language and communication preparation for emergencies and disaster management effectiveness, through the guise of whether translation is a 'broad concept of oral translation and

written translation' (pp.3). O'Brien et al.'s (2018), evaluation identified areas of weakness, and provided suggestions for improvement for each of the nation's assessed, which can motivate changes to improve the experiences of those affected by disasters and emergencies.

Availability can be considered as the minimum standard for information being retrievable in a specific language (Clarke & Isphording, 2015; O'Brien et al., 2018). The level of availability in this thesis will be separated into three categories; open, limited and closed. Open availability is when information is readily accessible without any additional efforts made by the requester, or any additional resources provided. Limited availability is the status quo of refugee management language use, as refugees need to put in formal requests for language support if they speak languages which are not part of the defacto language framework in use. Closed availability is when the information is not assessable or provided, and there is no effort being made to provide support.

Logistic and ethically dubious barriers to accessing information in a desired language include financially, where the requester is required to contribute to the cost of production (Kabranian-Melkonian, 2015), as well as practical, where the production of material takes time (O'Brien et al, 2018), and thus the requester needs to wait days to acquire information in a format that is equivalent to the information provided in other languages. Ethically, the situation of providing language support is complex when language rights, and access to language material is considered a human right and fundamental freedom. If language rights are considered as an absolute minimum, then anything less than open availability is unacceptable and a breach of

human rights, from an ethical perspective. Contrastively, if the position on language rights taken is that access is a guideline principle which should be aimed for, but is not a mandatory minimum, the status quo, of an ad-hoc compromise, is acceptable ethically. The UN's position on language rights is the foundation for language use as a human right, as well as the right to receive information in a language that is understandable (McAuliffe, 2009). Research across language policy and planning agrees overall, that language access equality, i.e., when all languages are available, is the standard that should be strived for, but implementing it is difficult and impractical at present (Clarke & Isphording, 2015; Broussard et al, 2019).

A significant issue to account for is when there is an imbalance between the language availability and accessibility, as differences can result in distrust and aggression towards the aid providers, as refugees could perceive the situation as discriminatory or favouritism (Kabranian-Melkonian,2015). When refugees feel that they are being disadvantaged because of their language, they are more likely to avoid accessing other support open-mindedly; this can escalate to distrust between the refugee population and the aid providers (Gerver, 2021a). The imbalance can explain, as an additional factor, the extreme situations that exist in the current Syrian aid camps, where terrorist and anti-aid provider sentiments have escalated to the point whereby the governance of the camps is under the forceful jurisdiction of terrorist organisations (ibid). Language access is not the only factor in how this status quo has appeared, but the imbalance in language support provided can be exploited to create social categories and identities where the refugee populations are one group, and the aid providers are another group, which are distinctive and separate,

which can be manipulated into an Us-Vs-Them mentality across the refugee populations (Gerver, 2021b).

When selecting a language to use in an emergency it is vital to consider ethical obligations, to avoid issues relating to the deployment of a policy into the unstable situations. It is unacceptable for both research and the deployment of human-based services to disregard any ethical issues (Broussard et al., 2019), as doing so can result in harm or injury to refugees or aid workers. A primary ethical issue to remember is that the population for disaster language policies will be refugees, who have experienced great instability; as their lives, livelihoods and social circles will have been in continuous upheaval due to the human disaster from which they fled. When thinking about language policy there are a few ways policy makers can pre-emptively account for issues, with the primary tool being the ability to continuously account issues which could affect the validity, impact or status of any policy deployed.

There are a few factors that can be considered when selecting the languages or varieties to use in emergency responses and refugee management, namely along the vectors of size, accessibility, and availability. In terms of size, languages need to have a significant population of speakers, both native speakers and second language learners, to ensure the use of the language would be able to connect the population with information (O'Brien et al., 2018; De Varennes, 2021). There is little agreement as to what counts as significant in terms of language policy and planning, this is the same in both general policy and in emergency policy (O'Brien et al., 2018), an issue that mirrors the question of what makes a language separate from a dialect. A

general rule of thumb is that a language policy in an emergency needs to be impactful enough in communicating with the affected population to justify the cost of producing the policy at the very least (Jacobsen & Landau, 2003). Furthermore, the speaker population needs to be within the affected area, as there is little point in having language resources for a population that does not understand the language; whilst this might seem like a simple nuance it is important to highlight and eliminate doubt as to the minimum requirement to always be accounted for.

In terms of accessibility, a language needs to be standardised, so that translation services produce preparatory resources (Austermuhl, 2014), such as information leaflets and direction signs, which can then be used to reduce the demand on the human resources, the workers, as refugees can be used to visually signpost rather than relying on ad hoc spoken directions (O'Brien et al, 2018). Standardisation is also required to be able to allow for the training and deployment of interpreters, with enough confidence that all interpreters can be used across the target area for the language populations in question (O'Brien, 2016), for instance if a policy uses a standard language, then it can be assumed that all speakers of the language will be able to, at least in part, understand the messaging from the interpreter. Thus, standardisation is a core factor that is needed to provide a stable foundation for the interpretive and translation en masse in emergency situations.

This subchapter highlighted how when providing language support for a target population, the rights of the individual should be maintained. Speakers should be provided sufficient opportunity to converse in a language they understand, and emergency policies should be designed with multilingualism as a core value.

Furthermore, the policies deployed to provide language support should adapt to cater for the demands and requirements of the target population. This thesis is investigating language support provided for refugees, and as such, the requirements and aspects of these migrants need to be accounted for. A refugee is classified as an individual who has flee from an at-risk region to a safer place, and in doing so they have experienced significant hardships. A common condition identified in refugees is Post Traumatic Stress Disorder (PTSD) (Johnsen et al., 2013; Bager et al, 2018), which developed from extraneous stress; and as such refugees, and their responses, are impacted by this condition. As such, the following sup-chapter will outline how stress is realised and how to account for stress in experimental research.

2.2.3 The Contextual Factor of Stress

As highlighted by Wright (2016), the contextual factors involved with the desired language use or planned usage requires the research to explore the additional factors that impact, or alter, the language realisation or processing. Originally, Wright's (2016) point was directed toward the sociolinguistics or historical elements within a society that direct or impact the acceptability of a language convention within that community. Examples include wars, conflicts or schisms based on cultural differences, where the use of terms identified with or used by the opposing sides may result in negative responses and attitudes towards an un-suspecting linguist. As this thesis was motivated by the language situation in refugee camps, where the population had fled due to human disasters, it is vital to explore the contextual differences between the refugee population and a non-refugeed group. In this case, it was identified that stress was a considerable difference between the two groups, and

as such, this subchapter will explore what stress is, both in terms of defining and measuring, as well as considering ethical methods to produce the contextual factor of being under stress for experimental usage as a constant environmental factor; so that representativeness in condition can be reflected in any study conducted in the thesis.

It is important to state here that the conditional factor of stress will not be the forefront aspect within this thesis. The replication of stress will function as to create a conditional simulated environment based on the real-world field work conditions which motived this thesis. Therefore, stress replication and induction will be the primary consideration guiding the following review and outlining of stress. An additional factor is that this that this thesis is not guided towards measuring stress responses, nor the response of the stress response. This is to say that the aim is to replicate a state of stress, and it will be within this continual state of stress that the effect of languages shall be tested and explored.

2.2.3.1 What is stress?

Stress is when tension is experienced, from which pressure is applied to the functional parts and elements of the body (Allen et al., 2014). The definition of stress itself is difficult, given that the term currently describes the concept of feeling pressure or tension, but the results can vary greatly depending on the type of tension, the location of tension & the root cause of the tension (Stall, 2004). Another consideration in defining stress is tolerance levels, which is the amount of tension a person can handle without being negatively impacted; within psychology this is called individual tolerance (Villada et al., 2016).

Furthermore, whilst defining stress itself as a concept is difficult, the responses of experiencing tension are easier to categorise and explain. There is a distinction between the length of time for which the response to stress occurs, which is directly related to the length and severity of the tension that caused the stress event (Kline, 1999). Another distinction is in the response itself being either positive or negative in relation to the physical, mental, or emotional status of the individual (Lupien et al, 2007). The following section will outline how the body reacts to tension and stress events, after which the distinction above will be explored.

2.2.3.2 How is stress realised?

The body responds to tension via the hypothalamus, which in turn orders the release of chemicals (also known as triggering) (Lupien et al, 2007). There are two pathways of response, the first being to activate the sympathetic nervous system which results in the release of Ephedrine and/ or Norepinephrine (Bong et al., 2016). The second pathway involves the activation of the hypothalamus pituitary adrenal axis, which commands the adrenal cortex and pituitary glands to release cortisol (Boucsein, 1992). Both pathways result in the same outcome, that of the fight or flight response (Broadbent, 1971). The fight or flight response is the body's natural way of dealing with the situation causing stress. It is the response of either physically/mentally attacking the stimuli to end the stress stimulant, or to run from the situation to alleviate the stress occurrence (Frisch et al., 2015). A third response is also theorised to exist and that is the Freeze response, which is where the body goes into a cognitive overload due to the stressing stimulus, whereby all reactions are frozen for a time (Johnsen et al., 2017). The strength of the response is proportional to the

pp. 62

hormone released, which is indicative of the tension experienced, lesser tension may not produce the extreme level of fight or flight, it will induce a moderate reaction which may not cause the individual to physically leave or physically fight, which of course is the extreme end of the reflex itself (Starcke & Brand, 2012).

Acute stress is when the tension, whether that be mental, physical, or emotional, is experienced for a short time and the stress response relief accounts for the entirety of the effects of the tension (Kline, 1999). This means that when a person is exposed to acute stress, they will have a triggered response which will dissipate rapidly when the tension experienced by the individual has stopped (Fanning & Gaba, 2007; Birkett, 2011). Examples of acute stress include calculating maths equations (Avancini & Szűcs, 2019) or quickly running down a street (Ashcraft, 2019). In contrast, when stress is longer term it becomes chronic, this is when the effects of the tension and subsequent biological response have been continuous and as such have negatively impacted the individual, either mentally, emotionally, or physiologically (Sandi, 2013). Chronic stress itself is produced by a compounding effect of multiple tensions experienced, therefore it is important to point out that chronic stress is a result of long-term acute level stress (Müller et al., 2009). With acute stress the response is negated rapidly after the tensions experienced and the trigger occurring, whereas with chronic stress individuals do not recover rapidly posts stress event and tension conclusion (Johnsen et al., 2017). To illustrate, being able to hold a glass of water in one's hand would produce tension in the muscles to hold onto both the glass and to retain the pose and structure of the arm and hand. A person could hold onto the glass for a short period of time without experiencing any

negative consequences as a direct result. However, if the glass was to be held for 12 hours consecutively, the holder would experience muscle fatigue and pain as a result. This is because the tension in the muscles start to cause damage and weaken due to continuous and prolonged experiences of tension. Therefore, the task itself begins with acute stress, however after continuous exposure to tension the task then becomes chronic in impact. It is possible to induce both types of stress, however it is ethically dubious to induce chronic stress in an individual, due to the nature that chronic stress may become irreversible (Pagani, 1989). For instance, a person can hold a glass of water for 12 hours, but it cannot be guaranteed that they will recover from the experiment as confidently as it cannot guaranteed that a person will recover in holding a glass of water for 10 minutes.

A further distinction of stress can be found in how a person processes or considers the tension that they are under (Quilici et al., 2005). If the tension does not exceed the tolerance level of the individual, in that the resources required to handle the demands of the tension are not exceeded, then the stress itself can be seen more as a challenge. When this occurs it is called eustress, which is when a person perceives the tension through a positive psychological state (Stall, 2004). In contrast, when a task or event produces tension which exceeds the tolerance level of the individual, when the resources required exceed that of the ones available, it is called distress (Hart, 2009). When this occurs, individuals can perceive the situation as either a threat or have a negative psychological response, which then changes the event or task, which preceded the stress response, to being seen as a danger or as a negative factor (ibid). This is important to distinguish as tension itself is not

necessarily an issue, but the perception of the risk and resources required to handle the tension is essential to understand when considering and understanding stress both experimentally and naturally.

Stress itself is a natural and normal phenomenon, as tension is required for all of life's functions (Arora et al., 2010; Allen et al., 2014). However, the extent to which humans overall experience stress is best explained on a continuum. At any one moment there will be highly stressed individuals and lightly stressed individuals, and this variation within a population is typical (McGrath, 1976; Santos et al, 2017). The high stress individuals could be those operating motor vehicles, tackling a difficult cognitive puzzle or processing complex information and data, whereas the low stressed individuals could be resting after a long day of work, or they could be falling asleep sat down on a chair. Throughout an individuals' day it is typical for their stress level to increase and decrease depending on time of day, the task being done and the regularity of the task (Pagani et al., 1989). What is atypical is when an individual, and the community, experience both a low level and a high level of stress simultaneously (Wetzel et al., 2010). Disaster events can cause significant and unexpected tension in A population, particularly if the risk of harm is increasing as time progresses exponentially or even proportionately (Warheit et al., 1996; Dückers, 2007).

When an event results in overwhelming levels of tension, it creates a deeper level of impact on an individual than chronic or acute stress can do. When this occurs, it is called trauma (Hart, 2009). The Yerke-Dodson Law (Teigen, 1994) can explain and illustrate how trauma develops, and the level of tension required in an individual to

become traumatised. The Yerke-Dodson law theorises that when a person is engaging with the task or difficulty, if the stress is low then the performance itself will also be low, creating a sense of calm in the individual (ibid). Furthermore, as stress increases, so does performance, up until a peak point, which is known as maximum challenge (Kline, 1999). The performance between the end of the calm state and the peak of the challenge is when an individual will be in a state of eustress, whereby they will be feeling the positive psychological effects of the challenge (Wetzel et al., 2010). However, if the stress increases further then the performance significantly reduces, and the individual experiences distress, which is the negative psychological realisation of the challenge (McGrath, 1976). The higher the stress level experienced the lower the performance becomes and at the extreme end the individual can experience breakdown or burnout which is when the individual feels exhausted and strained from the task itself (Allen et al., 2014). This can be used for a physical task, whereby the benefits of exercise eventually taper off until the muscles can no longer operate; it can also be applied to psychological tasks and experiences, whereby an individual will be unable to cope with any further information as a result of too much being forced upon them. This latter point is what refugees experience, as they will have experienced significant change in both lifestyle and situation of life to the point where they may develop stress disorders such as Post Traumatic Stress Disorder (PTSD) (Johnsen et al., 2013).

A further point to consider with refugees, is that whilst they are experiencing trauma because of them becoming displaced from their homes and communities, the difficulties they experience are not likely to decrease in the short-term following the

disaster event. When considering Maslow's hierarchy of needs, refugees will be focusing on a lower level of the hierarchy then they would have previously been focusing on prior to the disaster event (Fierros et al, 2017). For instance, shelter for refugees is not stable, or even sometimes a safe environment, and so the refugees will not be likely to relax in whatever environment they are in, so long as there are reminders that safety is not guaranteed. These reminders do not necessarily have to be physical reminders, in fact there is evidence to suggest that memories can trigger a stress response due to PTSD in refugees (Johnsen et al., 2013).

With refugees it is established across psychology that both the individual and the wider community experience chronic stress, via trauma (Bagilishya, 2000; Hart, 2009; Johnsen, 2013). Despite advancements in therapy, trauma is difficult to manage and can manifest differently within individuals (Hart, 2009). Furthermore, follow-up aftercare for trauma is expensive and difficult to provide en masse to larger sections of the population (Bager et al., 2018); as such it is an unfortunate reality that refugee populations are typically more likely to carry trauma without professional support.

At this stage, it is important to recognise the difficulties that researcher must account for when considering investigations into stressed contexts, particularly when the target population is impacted by and still experiencing chronic stress (Gerver, 2021b). It is dubious ethically to pursue any evaluation or experimentation on a refugee population, without taking steps to ensure that the refugees themselves have adequate professional care for their individual traumas experienced (Jacobsen et al., 2003). This is also difficult, because if a refugee is not supported, but the

experimenter implies that support can be given, then the refugee may engage with the research to obtain further support, rather than to fit the research characteristics, thus creating experimental bias (Kabrianian-Melkonian, 2015). The dubiousness of responses, particularly towards demographic information and questionnaires, is a factor that cannot be overcome easily with refugee populations. The relationship between the veracity of participants weakens as the living standards and status of the refugees reduces sharply, when the refugee is under the impression that they will receive any further support from the researchers. In sum, refugee populations are more likely to lie, or manipulate the truth, so that they can be part of the projects.

In short, refugees who are actively displaced will be experiencing significant challenges related to human survival and as such should not be subject to experimentation. This is physically the case for any experimentation involving or considering inducing any form of stress, as the likelihood that the individuals would be already under greater stress than average (Jacbosen & Landau, 2003). It is also important to note that as stress increases, it will only decrease if alleviation of stress is provided, however refugees are unlikely to be able to alleviate any form of induced stress in experimentation (Warheit et al., 1996). In fact, the majority of actions taken by refugees are under-duress, and that compounding variable is why it is ethically indefensible to promote or support experimentation of refugees. In research (acute) stress can be induced, provided that the stress is reduced shortly after the experimentation, and that the study will not cause any long-term damage to the participants (Allen, et al.,2014). Given that refugees are already stressed, it cannot

be said for certain the impact of experimenting will have on refugee participants (Kabrianian-Melkonian, 2015).

Stress is a response to an event or stimulus (Bong et al., 2016), as such several different methods have been developed which allow us to measure and categorise stress. There are two primary categories for measuring a stress response, the first being subjective and the second being objective (Villada et al., 2016; Allen et al., 2017). With subjective measurements, the aim is to assess the persons self-reported experiences of the tension and stress response, this can be done through questionnaires (Clover et al., 2020), such as the General Anxiety Disorder Assessment, or the Depression Anxiety Stress Scale. These questionnaires are self-reported, and the participants are asked to measure their experience on a set of pre-made scales, typically Likert in design (ibid). The results of these questionnaires can then be combined to create an overview of the general level of stress experienced across participant groups, as well as to build individual case studies. An issue with subjective measurements is the self-reported nature of the data, however, to counter this, standardisation of question strategy and question archetype is used.

In contrast, there is the second category of measuring which is objective. This is when data is collected based on physiological responses, with quantitative data collected (Smith et al., 2000). Examples of objective measuring of the stress response include heart rate, eye blinking rate and salivary cortisol rate (Villada et al., 2016). A stress response with both heart rate and eye blinking rate is indicated by an increase in the number of repetitions of the muscles involved, i.e.: the heart beats faster and the eye blinks at a greater frequency (Pagani et al., 1999). Both measures

require the pre-stress event levels and frequencies to be measured, so that a comparison can be made between resting state and the stress response state. The stress response both measures occurs rapidly after the neurological triggers are released (Ibid). Additionally, the recovery from the stress response is equally as short. In contrast, the concentration of cortisol in a person saliva can be an indication of stress response, provided that the concentration increases after a stress event (Abelson et al., 2014). The peak time to test saliva for a stress response is about 30 minutes after the predicted onset of the stress event and the triggering of the stress response itself (ibid). If there is an increase in the saliva cortisol concentration it indicates stress, furthermore the greater the increase the more likely that the impact event was a greater stress for the individual, as the stress response was greater (Lupien et al., 2007). The calculation of cortisol concentration requires enzymes to identify the biomarker of stress, whereas heart rate and blink rate are calculated by the sum of occurrences in a fixed time such as heart rate per minute and number of blinks per second.

2.2.3.3 How to induce stress in experimentation

A stress response can be induced for experimental purposes, with the aim of either measuring the individual response to stress or using stress as a method to replicate conditions which commonly produce a stress response (Sandi, 2013), in which case the aim is to measure an individuals' performance whilst undergoing a stress response. This thesis will fall into the latter category, as stress itself is not the primary consideration for investigation.

There are multiple ways to induce the stress response experimentally, which fall into three primary categories; firstly, there is the physiological stress which can be produced by getting individuals to perform physical tasks (Pagani et al., 1989). Previous tasks used include, experimenting on individuals during and following periods of exercise, such as running (Lupien, 2007). These tasks are categorised as requiring a muscular tension stress to be built up in the body, and that tension is what produces the stress response. The second category is emotional and social stress, which is when an individual is investigated whilst they are experiencing external mental pressure (Kiyonaga & Egner, 2012). Examples of this include the Trier-Tier Stress model (Birkett, 2011), which involves participants performing tasks in job interview style situations where they are being monitored throughout the tasks, without receiving feedback. This produces an environment where the participant feels they are being judged on their actions related to the task. This pressure creates a stress response based on the emotional pressure they feel in the environment. The third category is mental or cognitive stress, which is when a stress response is triggered whilst and individual is processing a logical issue or a mental concept (Sweller et al., 2011). Examples of cognitive stress include numerical equations (Imbo et al., 2007; Raghubar et al., 2010), memory recall testing (Lavie, 2010) and shape identification tasks (Garaven, 1998).

Using physiological tasks to elicit a stress response is based on the biomarkers of stress, emotional tasks are based on pressure from the sociological environment and cognitive tasks are based on internal pressure on the memory and processing potential of an individual brain (Todd & Marois, 2004). All three of the style of tasks

have been used in previous studies to elicit stress (Huttunen et al., 2011; Hartwright et al., 2018) and thus they can all be used as proxy measures for measuring stress, and for assuming a situation is stressful if the tasks have been deployed. Using cognitive measures is the easiest to deploy in online-based research, simply because the mechanism for inducing a stress response is found internally of an individual whereas both physical and sociological measures require a greater burden of observation to ensure that the stressed response is induced.

It is also important to remember that experimental stress must always be reversible, and as such acute stress is the only stress response that should be induced. This is because chronic stress can have long-term impact on an individual, whereas acute stress is repairable following the stress response and it can be elicited quicker with consent.

Cognitive Load Theory (CLT) (Sweller et al., 2011) specifies that there is a limited amount of processing power in the human brain, and that given sufficient stimulus and tasks, the total processing power can be overloaded resulting in the reduction of successful responses and actions towards the tasks themselves. Cognitive load theorises that if a person is under enough mental pressure and tension they will begin to perform less accurately in mental-based tasks and start to experience a greater error rate in their actions (ibid). CLT works in tandem with working memory theory (WM) (Engle et al., 1999) and the multi-store model of memory (Atkinson & Shiffrin, 1968). Working memory theory proposes that there is a specific maximum amount of capacity that the brain can work with, and that this processing element of the brain is constantly being refreshed to process all parts of human existence

pp. 72

(Kiyonaga & Egner, 2012). Multi-store model memory is a theory that highlights that there are multiple types of stimuli which are each processed in different manners in the brain, and to utilise the mind fully, a range of stimulus is required (Atkinson & Shiffrin, 1968; van Merriënboer et al., 2003). Examples of stimulus include audiological and visual (Klingner et al., 2011). The aim of applying cognitive load using CLT is to overload the amount of power that can be used in the working memory, by using a range of stimulus to overload the individual stores of memory. In doing so, cognitive load will reduce the memory capacity of an individual, making the production of responses to a task more difficult, particularly with longer-based tasks (Garaven, 1998; Kiyonaga & Egner, 2012). An advantage of testing using CLT is that the working memory resets regularly (Cowan, 2008), so overloading does not facilitate long-term damage to the brain.

CLT is an established feature in behavioural sciences, with the results being examined across multiple disciplines. For instance, in linguistics cognitive load has previously been used to investigate speech prosody in simulations (Huttunen et al., 2011), and in learning capacity for surgeons in simulation testing and assessment (Wetzal et al., 2010)

There are two types of cognitive load that can induce a stress response, they are intrinsic and extraneous (Conway et al., 2005). Intrinsic cognitive load is directly related to a task that participants are asked to perform or respond to, whereas the extraneous are the additional factors that are not directly related to the task response itself (Dowker, 2019). A simplification is that there is an intrinsic cognitive load as a

person calculates an equation, however if they are being watched, then the pressure of being watched is an extraneous new load that adds to the overall cognitive load.

Across psychology and behavioural sciences, a range of intrinsic loads have been produced and tested previously. These include using mathematics (Raghubar et al., 2010; Saeed & Sasangohar, 2017; Hartwright et al., 2018), and the internal processing of maths to produce a stress response, or using working memory recall (Imbo et al., 2007), such as getting participants to remember information prior to completing an additional task, and then asking the participants to recall the information they were asked to remember. Both designs highlight how intrinsic load is also linked to the difficulty of the task, where the quantity of resources required in the working memory can be varied to produce different levels of stress response.

Given that intrinsic load is directly related to the task which induces a stress response, it is easy to vary the stimulus to produce variable levels of cognitive load and cognitive stress, however with extraneous loads it is more difficult to control all factors that could produce a stress response. This is because the extraneous load is directly related to the non-task base elements, for which most of an environment can influence. Example designs for inducing extraneous cognitive load include; providing continuous feedback on accuracy for tasks (Dowker, 2019), whereby the participants are informed of whether not they are scoring favourably or negatively, or giving the participants the incorrect instructions on how to pass a task (Stall, 2004; Müller, 2009), this makes the participant start to second-guess the situation and in doing so forces them to divert processing power from the task at hand, therefore, reducing the working memory capacity during the task itself.

Given that the aim of cognitive loading in experimentation is to overload the working memory, it is common practise in CLT that a range of intrinsic and extrinsic stimulus and cognitive load are applied to ensure that a stress response is induced (Ashcraft, 2019). This is particularly important for proxy stressing, whereby the stress level of the individual themselves is not monitored or measured in an experiment. An example of using both styles of cognitive load can be found in using mathematics as the cognitive task in experimental conditions (Klingner et al, 2011). The calculating of an equation requires intrinsic cognitive load as the task itself is demanding on the working memory. There is an additional load applied specifically for mathematics which is channelled through maths anxiety (Hartwright et al., 2018); which is when a person believes that they are bad at maths and that previously they have felt pressure to perform highly with equations thus, attributes the situation to previous stress response inducing situations (Ashcraft, 2019; Dowker, 2019). Deploying a maths task can induce an intrinsic cognitive load, as there is a processing power for the task itself, whilst also inducing an extraneous cognitive load, which is the processing power diverted from the task to produce anxiety in the individual.

2.3 Language issues and features in disaster response

2.3.1 Current on-the-ground communication strategies used in emergencies

Clear communication and easy-to-deploy language strategies are vital in managing a population impacted by a disaster (Lachlan & Spence, 2009). If a community does not understand the messaging given by officials, the results can increase the risk level for the people; whereas, if the messaging can be accessed without issue, the

risk can then decrease (assuming compliance) (Sandman, 2012). As such the management of populations in a disaster event, or post event, are reliant on communication to produce the desired changes or effects. The movement of a group requires the direction of travel, the target or end point location, and the rationale for moving, all to be communicated from the organiser in a way in which the population will comply (Kuligowski & Omori, 2014). To facilitate this communication, management strategies need to consider the different methods of information delivery and the effectivity of the delivery mechanism. There are multiple methods for delivering information which, includes using modern technology such as SMS, as well as the more traditional, such as signposts (Wogaler et al., 2002). The effectivity of the mechanism can be measured according to the likelihood that the desired effect occurred (Carter, 2008), such as if the target is to get the population to relocate then you can measure the communication system to deliver the information on the relocation orders, simply based on the number of people that moved. In disaster management, the simplicity of the information is not necessarily an indicator that the communication strategies frameworks need to be equally simple (Arai, 2013b); rather inversely, with more care and consideration being placed on the information delivery as it is vital that information be received clearly and consistently throughout the disaster event (CDC, 2014). A major crux of emergency planning and disaster management is that the distribution of information and the effectivity of the information itself in deliverance is highly dependent on the level of preparation that has been given to communication for the specific event type prior to the onset of the disaster event (O'Brien, 2016).

The following sections will explore the difference between planned responses to a disaster event and the unplanned, following which will also be a consideration of the universal issues that are faced in terms of communication regardless of the level of planning taking pre-disaster. For clarity, the latter section will consider those known which are synonymous with the context and situation of the delivery information rather than the language use, or language used. The final section in this sub-chapter will be an exploration of the factors which are established as being impactful in disaster communication, and as such should be accounted for in planning policy for disasters.

2.3.1.1 How are language responses to disasters produced?

There are two levels of guidance which are deployed in tandem to organise and regulate the language use; and both levels are important to account for when considering language use in disaster and emergency management (O'Brien et al., 2018). The first level of guidance is set by a language policy, this includes guidelines as to which language to use; the topics and themes to be discuss and the minimum communicative requirements that need to be facilitated (Ammon, 2006). The second level of guidance set by language framework, which is the more in practice element of communication in disasters, these are the mechanisms and regulations that dictate the communication style, method & delivery approaches taken in a disaster (O'Brien et al., 2018). Language policies will consider issues such as who is to receive the communication support and why a language should be included and supported for use in a framework, whereas language frameworks consider issues such as what structure any messaging should have, the resource requirements and logistics of

deploying language support and the extent to which the minimum human rights requirements have been facilitated and fulfilled in deployment (Gomes, 2020; Leibowitz, 2015). For the avoidance of doubt, this thesis will consider language frameworks both preparatory and in use (past and current); however, the focus will be on improving language frameworks and policies in the preparatory phase rather than supporting or adjusting in-use policies. For clarification, this thesis will explore which languages should be used, A or B, rather than notions like will messaging be more effect when using X or Y.

2.3.1.2 Preparation

When analysing previous communication strategies in a disaster, it is important to register the probability of a disaster event occurring (Donner et al., 2007). Specifically relating to whether the disaster had a language policy that was prepared for the event or whether the policy was produced during or following the disaster event itself. If there is a pre-planned and actionable language policy prior to a disaster, then the techniques and features of the policy will be more effective, efficient, and targeted to the impacted population, than when the policies are produced in situ and ad-hoc (Manuel, 2014). When planning is made retrospective to the commencement of the disaster, the focuses can be unbalanced and short-sighted to long-term issues and difficulties (ibid). This distinction might seem extreme; however, it is important to remember that a disaster event stresses all elements of the civic and social infrastructure resulting in reduced effectiveness. The same is true for the language infrastructure, if there is no planned policy, then the ad-hoc becomes the primary policy, often using the logic of *if it worked once*, then it is often sufficient to continue

with (O'Brien et al., 2018). Whereas preplanning strategies and systems to support or facilitate communication during a disaster are more likely to be able to cater for a wider variety of individuals and language groups (Mayhorn & Mclaughlin, 2014).

There are two types of disaster response, depending on whether the disaster was expected to occur, or if it was unexpected and sudden (Keselman et al., 2005). Each situation has communication realised differently, and the techniques deployed are equally as diverse. There are techniques which overlap, and after time, responses to sudden disasters can stabilise to being that of standards found for expected disaster polices. The following section will outline the methods in which communication is realised in both situations, aiming to highlight the distinctions between prepared and ad-hoc communication frameworks and systems.

2.3.1.2.1 Expected disaster events

The key to a successful post-disaster management strategy is the infrastructure that is available for deployment. The infrastructure can vary greatly between the types of disaster event, and the aims of the response. For instance, an alert system for an incoming missile strike, which would notify the population of the risk, can help prevent individuals being situated in high-risk areas, such as under bridges (which can collapse due to the geological tremors and shock), during, and after the initial disaster event (Sutton et al., 2014). In contrast, a publishing house, or printing infrastructure would provide support for medical disaster events, such as an epidemic (Keselman et al., 2005; Williams et al., 2016), better than an alert system; as the aim of the management is to inform the population of detailed risk, such as hygiene

techniques or how to spot the symptoms of the disease, whereas the alert system gives short, simple information for a single risk. In expected cases, the deployment of strategies and infrastructure is dependent on the level of effort given pre-disaster (CDC, 2014). Within disaster management, the core caveat is the ability to deploy techniques as quickly as possible following the onset of the disaster; and the bulk of effort should be placed pre-emptively rather than reactively (Taibi & Ozolins, 2016).

From a language perspective, the expected disaster events would have language provisions that are appropriate to the communities impacted, in a format that can be accessed or understood. This guideline is referred to frequently by disaster policy researchers (O'Brien et al., 2018), as well as by political guidance (Jones & Askew, 2016). The delivery of language materials can take on a multitude of forms, from the printed written to the interactive services with interpreters (Tanner & Obrecht, 2015). The training of interpreters is time-consuming, and logistically more difficult to manage when compared to the physical written resources (O'Brien, 2016). In sum, multiple formats for each language selected would be essential for an effective premade disaster management plan; this would allow information to be distributed sufficiently both in the disaster event and in the recovery, whilst also providing a foundation for the interactive services for more complicated needs of refugees, such as further processing.

2.3.1.2.2 Unexpected disaster events

Unfortunately, disaster events can occur unexpectedly, and with that, populations surrounding the disaster event, and those impacted by the event itself are likely to face high levels of risk without disaster management techniques and strategies being

even considered prior to the event (Sandman, 2012). This issue is rooted in how difficult it is justifying disaster preparations as it is politically and fiscally difficult to provide evidence to support work on something which may never be used (Carter, 2008). Ultimately this leads to situations of disaster where little effort, or sometimes no effort, has been devoted to organising a response to a disaster (Staub, 2009). When this occurs, strategies need to be made ad hoc, in that they are made to meet the demands as they exist at the time, which might seem logical except for the fact that disasters are highly fluid and fluctuating situations, and to only be preparing for the current situation is a reactive management technique which does not counteract the next impact or effect (Carter, 2008). For instance, consider a disaster event that requires the population to be relocated, if the ad hoc system is in place, then the focus first will be moving the population. However, if there is little effort towards producing sufficient infrastructure to house the moved population, then they will arrive to unprepared environments (Koroutchev, 2017). As a result, the population that has been relocated will arrive to an area that has a lack of shelter, food, and medication, and such they will not begin to calm, relax or settle; predominantly because the environment does not have the capacity to provide those services (Gerver, 2021a). This then leaves the population in a state of flux.

Whereas a planned evacuation, would account for the resources required for the evacuation relocation points, such as, building shelters or stockpiling food, which would allow the population to begin to settle, even if it is only on a temporary basis. In summary, the unexpected disaster event is more likely to produce secondary side-effects from the disaster management techniques, because they are reactive rather

than proactive (Carter, 2008). Additionally, the same can be said for when a disaster is expected but the effort given to the management and development of techniques in advance prior to the disaster event is less than required (ibid). This is not to say that all disasters must be treated with equal effort, it is rather to highlight that appropriate planning is just as important as adequate awareness (CDC, 2014).

In terms of language use, the workers in the unexpected disaster events are forced to use whatever pre-existing policy of language exists in the community, to distribute communication and information (Ramirez et al., 2008). If there has been no planning related to language or disaster management for an event, because it was unexpected, then the emphasis will be focused on the community to provide the language framework in action (Taibi & Ozolins, 2016; TWB, 2017d). This is realised through the recruitment of bilingual speakers to act as de facto interpreters on a volunteer basis, on the argument that any communication is better than no communication (Morini, 2013). In the initial phases of a disaster response, the volunteer workforce can be a vital tool to facilitating management goals and reducing the risk-of-harm to a community; however, the longevity of the workforce is the fatal flaw of this approach (Maestri & Monforte, 2020). As time progresses, the volunteer workforce will weaken and become tired, and as such errors are more likely to occur or worst of all the volunteers will quit and the lack of language framework will be a dominant factor affecting the work of the governmental agencies or any operating NGOs. (ibid) The issues related to interpreters, as a workforce, will be revisited later in this sub-chapter, for now, the primary point is that there is operational bias towards relying on volunteer interpreters, specifically, viewing the use of such a workforce as

a long-term solution to the language issues faced in the field; an approach which is flawed.

2.3.1.3 Universal issues

Regardless of the level of expectedness, or the level of preparation that is taken towards addressing communication in post-disaster management, there are issues related to active communication on the ground that can occur. These issues are related to the errors or barriers that exist when communication occurs within the high-stress, or high-flux environment that are when disaster management techniques and strategies are at the most vital stage (Carter, 2008; TWB, 2017d). The largest cause for the communication issues in a disaster recovery strategy is time, or a lack there of (Ng &Hamby, 1997). Afterall, a post-disaster event, or post-commencement environment is unlikely to have a reduction of risk to the extent that the safety of those involved is high enough for slow communication. For instance, when processing the population following an explosion, the post-explosion environment carries itself unique risks, which were not directly connected to the risk from the explosion; as the impact of the high-pressure blast wave could damage infrastructure, and the damage done to said infrastructure could pose additional risks (Sandman, 1989; 1993). Consider a bridge, from which standing under it, during the explosion is high in risk, but the risk is intensified as the time increases post disaster, as the structure will weaken further due to gravity and the cumulative effects of the damage.

There are a range of issues which are specific to the environment surrounding communications and discourses, particularly the physical elements that can cause

follow-on effects to reduce the effectivity of communication between government agencies, interpreters, and the population (Wogalter et al., 2002; Carter 2008). When considering language and communication problems in disaster situations, it is best seen on a spectrum. With the extreme negative of the spectrum being that of a complete and total communication barrier, where communication cannot be facilitated because the language of the speaker does not correlate in a way that is understandable to the language of the listener (Lachlan & Spence, 2009). On the opposing side of the spectrum, is total open communication where there are no language barriers or communication issues occurring whatsoever. Many emergency responses and disaster planning, hinges on there being more of a swing on the spectrum towards total rather than no communication (Sandman, 2012), which is, in itself, slightly apprehensive in an assumption that communication can be facilitated and reached easily. This is not to say that communication cannot be reached in a disaster-based environment, it is a rather criticism of policymakers where the focus is on producing management plans with an assumption that communication is not something that requires detailed planning.

Commonly found issues in the field relating to communication, more importantly the breakdown thereof, often relates to the physical environment surrounding the discourse (Suied et al., 2008). This can include audiological issues, semantic issues, or mistranslation. The following section will highlight common issues of moderate significance that occur in the field of disaster management; with consideration as to why the problem exists and how it manifests from a linguistic perspective. The primary element considered is that of a discourse using an interpreter, as a facilitator

for communication. For instance, when an on-the-ground interpreter is considered, the demands are time-limited (Farazmand, 2001), as the interpreter does not have time to delay their work to ensure the conveyed messages are completely equivalent linguistically to the originally received message. The four common issues are as follows:

1) Semantic decay (Owen et al., 2009; McAuliffe, 2009) – this issue arises from the literal meaning of the original sentence being lost during the processing of the interpretation. This effect is linked to the brain incorrectly predicting the message being sent. The decay rate is proportional to the number of interpretations, or conversions, required; at a minimum if an interpreter is required for spoken discourse between two speakers of different languages, there is one conversion; thus, the decay rate is at the lowest point. However, if further interpretations are needed the decay rate increases. When there are multiple person's involved, the effect of semantic decay is known colloquially as the 'Chinese whispers effect'.

2) Decay resulting in pragmatic reliance (Bassnett, 2013; Morini, 2013; House, 2014) – This is when the sentence is mis-heard in that there are missed lexical items, or whole units, rather than singular units being lost. When a sentence is incorrectly heard, but the sentence is still linguistically intact, then additional, and potentially critical information is lost without there being any indication to the listener that a parsing repair is needed. When there are units lost that provide additional details for clarity, such as directions or additional attributive information, then the listener used their pragmatic and contextual information to add to the meaning of the sentence. Whilst there is nothing inherently wrong with using pragmatic context in most

discourses to aid communication understanding and processing; the lack of clarity within situations of higher risk or when the environmental ques are ambiguous then the reliance on pragmatics can lead to a negative outcome. For instance, if an adjective is lost when giving direction to go to the blue safe house in a residential area, then the listener is merely instructed to head to a house; which leaves the listener to rely on contextual ques; but the taken information is limited since there is an abundance of houses in the region; however, if the adjective was heard then the mission of reaching the safe house would be easier to comprehend since the environmental ques would be filtered by the adjective of blue.

3) Direct translations vs Get the gist (specific's lost, such as pronouns and adjectives) (Larkin et al., 2007; Taibi & Ozolins, 2016) – a well acknowledged issue in translation studies is that of the 'get the gist' approach to interpretation. This is when the interpreter converts the semantic information in a basic format, rather than converting all heard elements verbatim; often this approach is taken due to time restraints in that both speakers are engaging in an active and fast-paced discourse. Translation studies practitioners have used this approach as a method to distinguish between translation services and interpretive services, as translations focus on retaining the linguistic equivalence between languages, whereas interpreters aim to convert the semantic elements as a priority, with linguistic specifics being omitted. When considering language policy, long term approaches favour the use of translation and maintaining linguistic equivalence, as the aim is to improve the linguistic competency of overall populations or to retain the equal standing and usage of two or more languages. However, short term policies, favour the use of

interpreters, as the focus is on removing communication barriers in active and instantaneous discourses (Tannenbaum-Barauchi et al., 2017). Whilst the approach to convert semantic information is positive, as the outcome is to reduce the communitive issues found when people discourse in different languages which they are not familiar with or can understand. The issues lie when the semantic information is not enough (O'Brien, 2016), for instance, when interpreters are used in discourse where the information given by the speaker is essential, such as when giving evidence or details about events; in which critical details can be missed in the interpretation, since the importance of information is ranked by the interpreter, who knows little about the topic or the circumstances. For instance, if a refugee was asked about troop locations of their aggressors, it is up to the interpreter to decide what parts of the refugee's speech is important enough to relay to those asking; however, in this case a translation would be better as the full information would be given (O'Brien, 2016). But due to the restrictions in place contextually there is little chance that translation is possible; thus, this is an issue which must be accounted for when approaching language policy, as this is an example of the specific dilemmas faced in short term policies.

4) Audio Decay (Keysar et al., 2000; Kuperberg & Jaeger, 2016) – misheard linguistic elements, such as individual phones of a sentence, can result in the incorrect parsing by the listener. For instance, if there is a decayed phonetic unit, such as vowel in the nucleus of a syllable, then the listener could incorrectly predict the transmitted vowel when repairing the word in their mind. The impact of incorrect repairs can vary based on the terms that are decayed; with minimal pairs carrying the greatest impact, since

there is a single linguistic unit that separates the words. Examples of the changes are: in English – Run Vs Ran [tense difference], in Arabic, ʾaʿdaaʾ "enemies" Vs ʾaʿDaaʾ "members" [semantic difference leads to incorrect social grouping] as well as darb "path" Vs Darb "strike" [lexical difference resulting in different threat levels in directions], in French – char "tank" VS phare "beacon" [presented contextual difficulties]. The extent to which the decay or incorrect parsing can impact the communication between speakers is related to language variation, which will be discussed in a later sub chapter (see *2.3.3.2 Linguistic Variation*).

Variation is a standard feature of language evolution, and deviation from a standardised form of language is natural; however, this fact is not accounted for in language planning for emergencies, beyond being an after-thought (TWB, 2017a). This is realised in the field with communication difficulties between speakers of the same language (Bassouiney, 2009, 2014; Horest & Cotter 2016). The differences reduce the effectiveness of the communication between interpreters and those in need, which is not in itself an issue of policy makers, as variation is natural. However, the issue is that the variation has been either ignored (Jones & Askew, 2016) or dis-regarded as a factor to consider by the policy planners. When the variation is not countered, it has resulted in the difficult situations that TWB (2017a) identified in the migrant crisis.

The NGO's and governments use interpreters to counter act the issues faced by language difference, however this is not a clear-cut solution as the refugees are still experiencing issues related to language (Federici, 2016; TWB, 2017b, 2017d). The first issue is that of resources, as there are not enough interpreters to meet the

demand in the specific places; there are two critical reasons behind the shortages.

Firstly, there are minority communities that have fled (Doğan, 2016), which can bring

along minority languages, which may not be standardised or even recognised by the

speaker's nation of origin. Finding a bilingual interpreter is difficult; as the population

of native speakers is limited, and as such relying on the chance that enough native

speakers of the minority languages relocated as part of the refugee community is ill-

advised, as there are no guarantees. The second reason is related to the fluidic

nature of refugee movement, as the phenomena is unstable and unpredictable

(Cadwell, 2021). Groups of refugees appear rapidly, for instance in 2011 there were

less than 10,000 refugees on the southern coasts of Europe; however, by 2014 there

were over 2,000,000 refugees (TWB, 2017a, 2017c). Currently, no nation has

developed infrastructure techniques to facilitate the almost exponential growth in

population that is caused by refugees and mass migration (Koroutchev, 2017). The

resulting effects of the lack of preparation and reaction efficiency range from slum-

like conditions of living to ghetto-like communities (Geiger & Pécoud, 2010).

The latter is of greater interest to linguists, as the closed off communities' exhibit similar

characteristics to ghettos from history, that are built not on social class or race but on country

of origin. A primary method for distinguishing between the separate groups of origins is via

linguistic characteristics, in particular: accent or phonological inventory (Hachimi, 2013;

Albirini, 2016). As when a person arrives in the aid camp and is introduced to the community,

the members of the community are analysing them to identify where the new person is from,

so when the new person speaks, the linguistic characteristics that they employ are a primary

method to define which group the speaker belongs in (Sankoff, 2016).

2.3.1.4 Abstand and Ausbau

A core question in linguistics is what defines a language as a single unique entity. Kloss (1967) proposed a dual classification model to categorise language, as a linguistic system, by two factors: the sociological context (Ausbau) and the linguistic distance (Abstand). Ausbau (roughly translated as by expansion from German), is when a language is defined according to socio-political differences and similarities, such as officialdom, status, and prestige. Whereas Abstand (language by distance) is when language is defined by the linguistic features shared. Kloss (ibid) proposed that language by Ausbau is derived from written languages, as conservative and prescriptivist measures are required for formal standardisation, which is the bedrock for the standardising of a language. In contrast, Kloss (ibid) suggests that Abstand is derived from spoken language, where the vernacular systems experience linguistic change and evolution naturally. Ausbau languages can be measured by social and political attitudes (Tosco, 2008); for if a group of people believe that they have a unique language that is connected to their culture exclusively, then they can declare it a language in its own right; regardless of the linguistic similarities with another language.

Abstand is less defined, than Ausbau, as there is little agreement as to the criteria that should be used to define a variety as being suitably different or linguistically distanced from another, enough so, to be classified as a separate language. There are multiple different approaches for the criteria for defining Abstand that have been presented since Kloss's (1967) original proposition, such as mutual intelligibility, lexical distance, historic genealogy, or phonological similarities (Phillipson, 2006;

Juffermans, 2015). An overarching issue with the methods presented thus far, is that the criterions lack sufficient rigor or justification as to the threshold point where a variety can be promoted to being considered a language. Even Kloss's (1967) original piece failed to address this issue, by opting to 'not now dwell on the problem of what criteria the linguists apply in measuring the intrinsic distance between language' (pp.30).

In his original proposition of the two-way categorisation of language, Kloss (1967) highlighted that Ausbau was the cultural and global norm for language defining; borne from the direct relationship between national power, identify and language; whereas Abstand was an applied, and defined, area. In fact, Kloss (1967) recanted that "'Abstand requires additional future work... in order to put the model into application" (pp..56); a statement highlighting the difficulty that linguists face when defining language by the linguistic properties; specifically, that the tools for doing so were non-existent at the time of proposition. Alas, despite Kloss's proposal being over 50 years old, the tools for defining language by Abstand are still in-development, and there is no accepted standard between academics, let alone nation states. This is resulted in the complex situations described in TWB (2017a), as well as in a plethora of other regions and subjects (for language issues in translation studies see Bassnett (2013) and Austermuhl (2014); for issues in psychology studies see Saiegh-Haddad (2004, 2014) Saiegh-Haddad et al., (2011) and Akbacak (2011); and for issues with language barriers see Zaiden & Callison-Burch (2012, 2014)).

The current situation is that language is almost exclusively defined by Ausbau approaches (Manfredi & Tosco, 2018), yet language is deployed, and thus the

language differences (or distances) are impacting the communication overall. Whilst at first glance, a solution to this is clear, hold two separate definitions for a language, whereby there is a defined marking between language by Abstand (which would be used by linguistics and communication specialists and in the field) and language by Ausbau (which would be used by political and cultural subjects), however in practise this is deceivingly difficult. In practise, an Ausbau language is typically the language variety in popular usage (Tamburelli, 2021) within the political or economic contexts, such as in governmental business or financial trading; and as a result, the specific variety would be presented as the correct or appropriate language to represent the whole nation. It is important to clarify that Ausbau languages, are in some-form, rooted in an existing language variety (Kloss, 1967; Tosco, 2005); whereby an Ausbau language is a language variety, but it is not the only form of language in the region from which it originates. Whereas Abstand languages are any language variety that is distinctive and distanced sufficiently from another to be classed as unique (Tamburelli, 2014).

When applying Ausbau and Abstand in practise, there are issues which are currently un-resolved, that of balance. Language is integrally linked to culture, national identity, and an individual senses of belonging, and in the event of a dual category system (as proposed by Kloss, 1967), there is a chance that the language of a nation could be separated into several different languages. This in theory should not negatively impact the national language or undermine the Ausbau language; since Kloss's model clearly states that there can be multiple Abstand languages which can be combined into a single Ausbau language. However, in practise it is unlikely that the

socio-cultural elements of the Ausbau language, or the Abstand language, which is

the predominant language politically, would not be impacted. For clarity, consider a

nation with a single national language (A), which is made up of four different

languages by Abstand (B,C,D & E), the national language is based on language B;

since it is the language used in the political sphere and within the capital of the nation

(see figure 2).
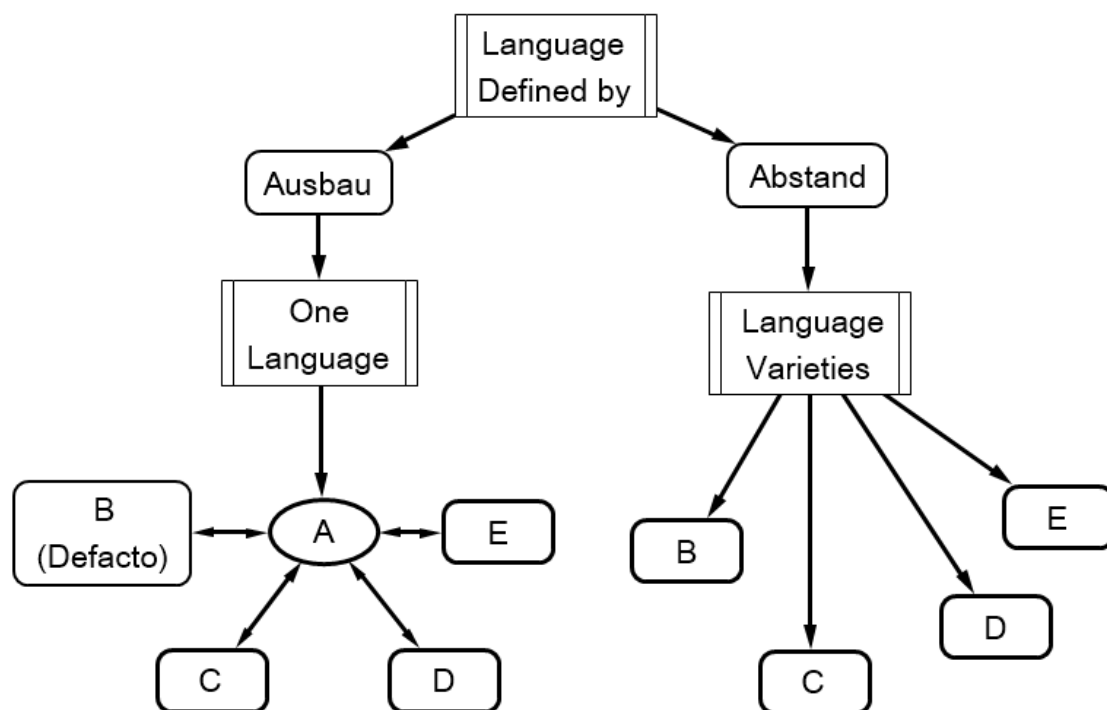
Figure 2: Visualisation of language demarcation criteria using Kloss's (1967) Model.

If the nation reordered the national language profile in line with Kloss's two system

model, then the language communities of languages C, D and E would gain

sociolinguistics prestige and power, at the expense of language B, since language B

would stop being the overarching and singular language; and thus, it was holding all

the sociolinguistic prestige of the nation. Despite the distinction presented by Kloss

(1967) being designed to aid linguistic research and language science, there is a concern (Trudgill, 1992; Tosco, 2011; Wright, 2014; Tamburelli, 2014) that the wide-spread deployment of this modelling would have significant impact on the language profiles, national identities, and sociolinguistic status quos. Ironically, in the event of balancing the definitions, i.e., Ausbau language definition is equal to that of Abstand, the issue would be that Ausbau-status languages would become reduced in power, and thus the established powers across the world do not often support any Abstand-focused approaches to gain popularity, since it could interfere with the political makeup of the jurisdictions and regions that they represent or govern. As such, a rebalancing of language definitions, and most importantly, language status is difficult to implement globally; as the re-balance would have far-reaching impact, from changed language status's to redefining bilingualism.

When considering bilingualism, the re-balancing of language definitions may impact the individual language user, particularly as the number of bilinguals may in fact increase. Consider this situation: there is a speaker who believes themselves a mono-lingual speaker of their national language; but when their language is assessed using Abstand, there are in fact 10 different language varieties that make up the national language. The speaker uses two of the language varieties in daily usage; would this speaker be classified as a bilingual speaker now? If so, how would the change of linguistic identity affect the speaker; they could accept the technical differences (the Abstand defined difference) without issue; but that is unlikely given the connection between language, identity, and nationality.

A secondary issue that has appeared in prevalence is that of bi-dialectialism, which is when a speaker can communicate in two different forms of language from the same language (Spolsky, 2011). The differences between these 'dialects' can be sufficient that speakers from within the same language cannot understand one-another. The notion of bi-dialectialism seems rather similar to the issue of Ausbau and Abstand imbalance (Kloss, 1967); given that there is a language, as proscribed by socio-political elements, from within which there are forms and usage that contain noticeable differences from others within the language, as explored as linguistic differences.

Additionally, it is important to clarify that not all languages by Ausbau are comprised of a set of languages by Abstand. An example of a situation of an Ausbau language which is singular by Abstand can be found within Eastern Europe, with Bosnian, Croatian and Serbian (Kapović, 2011). These three languages are stable languages by Ausbau, with each having national status and protections; however, by Abstand they can be considered close, potentially close enough to be a single language (ibid). To redefine the three languages into a single language could impact the social status's of the nations, which historically have been tense and war-fuelled (Massey, 2021), and thus to carte-blanche redefine there would be countenance to the peace processes occurring (ibid). When assessing whether a language is intelligible an Abstand approach to defining language would have been required, given that the question of understanding is rooted in the linguistic aspects and features singularly.

Internationally, regarding language policy, the opposition to redefining a language is connected heavily to the diplomatic language policies, in that any attempt to reduce

one of the core top six diplomatic linkages of the world is met with opposition not only from the speaking population but also from the official statuses and governments (Laverack, 2015). Whilst Europe was described as Ausbaucentric (Tosco, 2008; Tamburelli, 2014; Tosco & Tamburelli, 2021), this thesis would suggest that Ausbaucentrism is also at the heart of the United Nations language policy as the socio-political elements and factors are considered the only consistently suitable metric to define a language in the United Nations; whereas policy created by linguistic distance would undermine the unified political approach that the diplomatic language policy is designed to produce. In some instances, diplomatic language policies are resistant to change, and particularly resistant to redefining what constitutes a language if it would reduce the size and power perceived by the language community; for example, if a language would be to split into two languages by the application of standard methods it could arguably halve the collective power of the language communities and governments representing them on the diplomatic global stage.

In sum, the re-balancing of languages and what comprises a language needs to be done with tact and care, as changes have the potential to have wider impact that simple classifications. These changes would be drastic because of the focus that language by Ausbau has had on the linguistic landscape of the globe. It is common for a language to be defined by Ausbau in every nation (Tosco, 2008); with the defacto method for defining a language being Ausbau. Tamburelli (2014) described the global socio-political system of Europe as Ausbaucentric, an apt term, given the central tenants of European language policy seem to favour Ausbau languages

rather than Abstand. Which explains how Europe's multilingualism is dependent on the demarcation criteria, such as Italy is either a monolingual state (by Ausbau) or a multilingual state (by Abstand) (Tamburelli, 2014).

Prior to any continuation of the literature, it is important to address a terminology issue that has arisen thereof; that of what is *language*. As explored above, there are two separate definitions of linguistic systems that are attributed to the single term, which can result in conflation of the two distinctions (which undermines Kloss's (1967) original point that there needs to be a split between the socio-political elements and the linguistic features). To address this issue, the term language will be used to define language by Ausbau, whereas language by Abstand will be defined as a language variety. This distinction is motivated from Trudgill's (1992) work on unifying terminology to distinguish between Ausbau and Abstand languages. Using language variety as the primary term for Abstand, implies the existence of linguistic variation, which is a foundation of language by Abstand. The next section will explore the issues facing language by Abstand, and defining language varieties in the field, given that there were Abstand rooted issues identified in TWB (2017a), which cannot be explained using Ausbau method (and it was vital to explore the issue of language itself, prior to exploring the issue of understanding and comprehension).

Ferguson (1959) proposed the term diglossia to describe situations where 'two or more varieties of the same language used by some speakers under different conditions' (pp.325). In a diglossic community the language varieties in use are categorised as either being a higher variety (H) or a lower variety (L), dependent on the socio-economic and political status of the variety and native speaking population.

In terms of social power, the H variety is the variety with the greatest prestige of those in the community, this can be realised as the variety being the medium of communication for the court, education, or government. The L variety, however, is commonly attributed to the language use in the household and for informal discourses. Ferguson (1959) presented Arabic as a language that is heavily and stably diglossic, in that there is the H variety of classical Arabic and the L varieties of localised colloquial Arabics. Arabic has since evolved further, with MSA replacing CA as the SA of the Arab-regions; with CA being limited to religious contexts and practises (Versteegh, 2014). The diglossic situation has also evolved, with MSA replacing CA as the de facto H variety (Ferguson, 1991), however, CA has not been relegated to a L variety; as the social prestige is still high with the variety (Saiegh-Haddad, 2003). There are now two H varieties within the Arabophone, as well as a multitude of L varieties; as such the situation has grown more complex with evolution without reducing the stability of the diglossic situation.

Ferguson (1959) presented a list of key features distinguishing between the H and L varieties in a community, which include the stability and standardisation; two features that are important to consider in disaster communication and refugee management. In emergencies, if the situation is limited to a single community, then the population will be relatively stable, and thus it can be assumed that the diglossic situation existing prior to the emergency could still be in force, examples for small scale single community emergencies include locally based fires in residential centres. However, if the emergency is greater in size than a local level and affects multiple communities then the population in question is likely to not continue any pre-existing diglossic

situations or status quo's; because of the population migration/relocation merging multiple communities and multiple diglossic situations together (Walters, 1996). Examples of larger affected emergencies include the Southern Refugee Europe Crisis as the populations came from multiple nations as well as communities, and thus the language landscape was different to that of the origins for the populations (Gardner-Chloros et al., 2016).

The second key feature of diglossia (Ferguson, 1959, 1991) is standardisation, which is equally important to consider, in disaster communication in emergencies, as translation services and interpretations can only be functional to a high level of accuracy if there is a consistent system to use for both varieties. So, in the instance of the diglossic community with both H variety and L variety, Ferguson's (1959) notion, that the H variety is typically more standardised than L variety, could result in the minority communities, associated with the L variety, being more disadvantaged than those who can communicate in the H variety; and thus, sections of the population might be treated differently, which can be an affront to human rights of the affected populations in the disaster. In addition, it is important to note that the diglossic community can have both an H rating and L varieties which are all standardised in some part, in which case the argument of imbalance of treatments and the human rights of the population would be satisfied as being non-existent in theory, however, in practice they may be bias towards the H variety because of pre-existing diglossic attitudes being pushed through or retained by organisations offering translation or interpretive services.

An expansion of diglossia was proposed by Fishman (1967), who suggested that the situation was more complicated than Ferguson (1959) originally stated. Fishman's (1967) core argument was that there can be a relationship between two or more languages, or varieties of the single language, that are not genetically connected i.e., there is not a pre-existing linguistic relationship between the two. Fishman (1967) proposed that the situation of two or more languages or varieties being used in a single community should be segmented into the sociolinguistic elements and the psychological elements. The sociolinguistics includes the status and perception of the language variety, whereas psychological elements would include the ability to use one or more of the languages and the ability to functionally differentiate between the languages of varieties in use. The two elements will be used in a spectrum and could exist concurrently in the same community without issue (see table 2).

Table 2: Fishman's classification of bilingualism and diglossia

		Diglossia	
		With	Without
Bilingualism	With	Bilingualism with diglossia	Bilingualism without diglossia
	Without	Diglossia without bilingualism	No bilingualism or diglossia

Ferguson's (1959, 1991) diglossia assumes that the population in question all speak the same language, as diglossia is within language variation. When considering the current status quo of South Europe and the refugee populations within the aid camps

there, it is important to recognise the limitations of Ferguson's diglossia as camps are widely accepted as multilingual not monolingual (TWB, 2017a, 2017d). Furthermore, when attributing for Fishman's concept of bilingualism, with or without diglossia, the situation becomes ever more complex, and thus language policy needs to account for multiple levels of variation, status, and intelligibility. A similarity with both Ferguson's (1959, 1991) diglossia and Fishman's (1961) bilingualism is that both language varieties in a situation exist with a two-way distinction between higher varieties, which are formal and official in use; and lower varieties which are informal and are commonplace in the home. As such language policy for monitoring and processing refugees needs to, at the initial stage of planning, identify the target environment for implementation. The key for policymakers, is what context should the information will be delivered in, as this may influence language selection (Sandman, 2012) i.e. through government notices or media releases then the H variety would be more appropriate from a sociolinguistic standpoint (Powell, 2016; Ibrahim, 2021), however if the information is to be delivered in an ad hoc informal discourse style then the L varieties would be more appropriate in this might engage with the community more so than the H variety, due to a sense of familiarity (Gass & Varonis, 1984).

Stępkawska (2012), highlighted that there is a critical difference between diglossia and bilinguals as proposed and used by both Fishman (1967) and Ferguson (1959); in that both a measuring different factors. Diglossia is a social phenomenon, and the measuring of it requires assessing the language use of a community overall, whereas bilingualism is a proficiency measure where individuals can be assessed for their

degree of bilingualism. An individual cannot be assessed for diglossia in the same way that bilingualism can be tested. Bilinguals can be measured as balanced or/and unbalanced (Fishman, 1967; Kremin & Byers-Heinlein, 2021), when considering the social power and status of either language spoken, whereas, both models of diglossia (Ferguson, 1959, 1991; Stępkawska , 2016; Al Suwaiyan, 2018) require the status quo to be unbalanced between the language varieties in the same community. With these core differences between diglossia and bilingualism in mind; diglossia is often applied in an intra-language context, whereas bilingualism is exclusively inter-language; so, an underlying issue is the status of what is a language. It is possible to have both diglossia and bilingualism in the same context, simultaneously, if both Ausbau and Abstand are deployed in equal measure. There is a possibility to classify a single community as diglossic by Ausbau, whilst also having bilingualism as defined by Abstand. This nuanced approach adds an additional level to the situation, but it splits the social phenomena from the individual linguistic proficiency; providing that the population in question accepts that a language variety can be a dialect and a language, and vice versa without reduction of social prestige and attitudes. This isn't to say that change will not occur, in fact it highlights how the stability, in a diglossic bilingual community, can exist.

2.3.1.5 Major issues

Considering the severity of disaster miscommunications, it is vital to consider what happens in the event that there are no appropriate or actionable policies established. When communication has broken down to the point where language is not available, the fieldworkers and populations impacted by the disaster event resort to

paralinguistic strategies to convey information (Carter, 2008; TWB, 2017a). Whilst this may seem a sidenote, given that the thesis is aiming to investigate wider language use than ad-hoc strategies towards communication, it is important to take stock of the status quo of language use and communication in the field. When communication has been a barrier, the resulting communication strategies are negligible and insufficient at facilitating basic information, this is particularly impactful in the processing of refugees, and the following section outlines examples reported from fieldwork cases of in situ communication strategies when there is no language connection or facilitation between aid workers, government workers and refugees.

The first example is the use of a map and pointing on it as a mechanism to identify the region of origin for refugee (Koroutchev, 2017; TWB, 2017a), which is a strategy that assumes the refugee will understand the task at hand and will also produce a true answer. The second example is the retraction back to para-linguistic gestures as a solo method of communication, which can be useful if the aim is to simply relocate a large group of individuals, however it is a not suitable for the transfer of information beyond directions or highlighting geographical information, such as 'look over there', or [pointing in the direction to travel]. Both examples shown highlight the reliance on non-linguistic methods that communicators are reduced to when there are language barriers, and this para-linguistic reliance reduces the effectivity of communication, as well as limits the depth of knowledge that can be transferred between the discoursers, such as between the aid workers and the refugees (TWB, 2017c).

There are two more examples of ad-hoc strategies that will be considered, both of which are related to the delivery of spoken discourse, as opposed to the previous

paralinguistic examples shown above. The first linguistic ad-hoc approach is when a language barrier is experienced, the speakers increase their volume of speech production as a counter for the lack of intelligibility (Lam & Tjaden, 2013). Whilst the use of volume in situations of unintelligibility or language barrier, there is an emotional factor in the rational for deploying this technique, particularly when it is performed by the refugeed, or displaced. When realising negative emotional states, such as anger or frustration, the use of speech volume as an emotive release is commonplace (Hart, 2009), and thus, when a barrier to communication is experienced those who are already stressed, or frustrated, are more likely to react in an emotive way; of which getting louder is a method (Rajan, 2019).

The final example is when the speaker adapts their linguistic choices, in terms of production, to improve intelligibility by elongating the length of production for each vowel (Lam & Tjaden, 2013). This is known as the Lombard effect, where an individual changes their vowel patterns, particularly in terms on exaggerated lengthening, to improve the intelligibility of their speech (Ngo et al., 2020). The impact of the Lombard in improving intelligibility in noise and context is disputed (Clopper & Bradlow, 2008; Rogers et al., 2010; Anand & Stepp, 2015; McLoed, 2020). Whilst this technique may improve the communication between individuals, it can also impair, particularly if the deliverance of the exaggerating technique is considered negative by the listener; in which they may ignore the speaker, or worse, expel frustration towards the speaker (such as by raising their voice) (Sandman, 2012). Regardless, the nature of using a disputed technique in a high-pressure environment where volatile and close-to-the-surface emotions are also experienced, is ill-advised,

as the repercussions of failing in communication is unacceptable (especially from a human rights perspective with the notion of equality and accessibility).

This section briefly outlined several examples of poor communication strategies that are realised in ad-hoc or under pressure disaster management environments; these highlighted the extent to which accessible communication is required, as shown by the fact that even when language barriers exist, those involved cannot simply disregard the discourse, or ignore the issue. When faced with the challenge of language barrier, the in-the-field operatives must continue, and that is why it is essential for emergency language policies, and strategies to retain a mindfulness toward the reality that is experienced and the importantce that effectivity be at the forefront of policy and planning for language use in disaster management and emergencies.

2.3.1.6 Planning language workforce

So far, the exploration of language in the field has focused on the realization of language in disaster events and management, with little attention to the advice or guidance provided for emergency communication in general. This section will explore the elements of communication that have been identified as either useful to facilitating compliance with communication (Sandman, 2012), or the opposite, i.e., elements that can hinder the communication overall. There are two critical parts of communication that were identified from the literature, the first being the role of the interpreter and volunteers in the field, who function as the workforce for during-emergency communication (O'Brien, 2016). The second part is the structure,

features, and style of the messages themselves, which is particularly important when preparing for emergencies.

2.3.1.6.1 Interpreters

The human workforce is a backbone to emergency communications in a disaster (O'Brien, 2016). From a communication standpoint, the human workforces are the translators and interpreters that are present in the situation during the disaster event (Cadwell, 2021). In this thesis, interpreters are considered distinctively separate from translators, with translation referring to the written and formal conversion of information between two or more languages, whereas interpretation services refer specifically to short-term re-iteration of a message between two or more languages. It is also important to note that translators focus on relaying the messages across languages without losing any nuances, such as style or pragmatic sub-text, between the languages, whereas an interpreter focus is on relaying the key information, regardless to the style of the original message (Doğan, 2007).

In the field, interpreters are vital to providing communication between the majority language communities, and the minorities (Taibi & Ozolins, 2016), whereas translators are vital for the development of physical language records and resources (Lindell & Perry, 2007). Interpreters are short-term workers, who help the management of the disaster in the immediate time during and following a disaster event, whereas translators are needed to develop longer-term responses. To explore the vitality of interpreters, a brief overview of the functionality of this workforce will be presented. Following this overview, a review of the strengths and weaknesses of using interpreters as a primary language policy tool, including the attrition effect on

the workforce, will be presented. The section will present a practical overview of using interpreters as a resource in disaster management language policy, to understand the delivery system of any language-based communication.

In emergencies, interpreters function as communication conduits when there are language barriers between two or more speakers (Federici, 2016). In multilingual discourses, the interpreter listens to the words spoken in one language and then relays the message in another language.

It is important here to remember the difference between interpretation and translation. The former focuses on converting the core information from one language into another (Doğan, 2016), whereas the latter aims to replicate the original message's features into another language with maximum equivalency (Bassnett, 2013). Given the emergency nature of disaster, this study will favour both equally, with the hope that the future policies will have the capacity to translate more than interpret.

When in an emergency, interpreters are used to inform minority language communities, whether that be to identify the risk, direct the evacuations or to process individuals in relocation shelters. The use of interpreters in emergencies is standard practise (Doğan, 2007), as they can perform basic essential communication facilitation with little cost, effort, or notice. Interpreters can also relay information on an ad-hoc basis, which is reliably fast in the field when aid workers or government agencies are working with the refugeed, vulnerable, or injured, as the interpreter can change the messaging to adapt to the individual's circumstances.

Within interpretation services used in the field, there are two different kinds of service, based on the recruitment of the workforce itself, there are the paid professional workforces and then there are the volunteers (Maestri & Monforte, 2020). The former is the better standard in terms of accurate and reliable communication, whereas the latter is the fastest to recruit and the cheapest to run, by virtue of being free to run. The deployment of a paid interpreter workforce requires adequate planning for the language support prior to the emergency event itself. When human rights are considered, and the goal of having equal language support for all in an emergency, then planning needs to account for each language in the effected region, before the event starts; a costly and hard-to-predict exercise, which has little guarantee of being needed within the workforces lifetime (Garnett & Moore, 2010).

Furthermore, as seen in section 2.3.1.2.2, many interpreters that are used in a disaster situation are volunteers (TWB, 2014); who are recruited ad-hoc for the role, in some cases they are nigh-conscripted by agency workers to resolve issues between the agency staff and the displaced, before the situation results in conflict or argument (Cadwell, 2021). Volunteers, as a temporary workforce is not detrimental to the facilitation of communication in the emergencies, however, if the workforce is not replenished with new volunteers, or paid staff, then the interpreter services will decline. This is commonly due to workforce fatigue and loss of volunteers by attrition, as the individuals leave the service or relocate themselves (Maestri & Monforte, 2020). Economically, it is beneficial for the government agencies and NGO's to have access to a cheap-to-run service, which volunteers certainly are, however each of the

volunteers still have economic and social needs which cannot be met by working without pay (Federici, 2016; TWB, 2017a). When the volunteers are working, the lack of readmittance itself is unlikely to reduce the standard of the service provided, with service continuing immaterial; however, the vitality of the services overall is at risk. There are little guarantees that NGOs can assume when using a volunteer workforce, as the interpreters could stop reporting for duty without any consequence to themselves.

When considering language policy for emergencies, it is vital that polices account for the economic situation of the workforce that can be used to relay information. If the approaches rely on volunteers in the community to interpret for relief workers or the governmental agencies, then the longevity of the policy can be questioned; whereas policies that rely on paid trained interpreters will face difficulties in both costing the service and the logistics of having sufficient staff for each occasion. Ultimately, the current approaches to emergencies, is to use a combination of both paid and unpaid workforces (Carter, 2008; Federici, 2016).

There are number of strengths and advantages of using interpreters in disaster management, namely that the workforce is adaptive, culturally proficient, and potentially conflict resolvers (TWB, 2017e). Interpreters are adaptive, in the sense that the messaging can be changed dependent on the individual's in the discourse, unlike premade resources which repeat the same static message regardless of the cliental. Interpreters can be recruited from the target communities for communicating with, in such cases, the interpreters will have a greater understanding of the socio-cultural norms that exist in the minority community in question (O'Brien et al., 2018).

This cultural awareness removed the chance that faux-pas's will occur, as the interpreter could counter-act any language which could cause upset or offence. Whereas the usage of an outsider (in terms of the community) does not offer such protections, as such, the background of the interpreters is a factor for policy makers to account for. Additionally, the use of humans as the communication facilitators allows for the frustrations from both sides of any language barrier, to be alleviated in a natural and social interaction (Bassnett, 2013). The de-escalation using a human actor has been identified as a mechanism for improving compliance with emergency orders (Sandman, 2012) as well as supporting the community's processing of traumatic situations (Hart, 2009). Therefore, using human interpreters is favourable to using machine or technological alternatives when the situational management has resulted in underlying tension or frustrations between the agencies and those impacted, particularly when a language barrier is one of the factors causing or escalating the discontent between both parties.

There are also weaknesses in using interpreters, such as there can be exploitive practises used (both on and by the workforce) (Kabranian-Melkonian, 2015), veracity can be questioned and the underlying issue of whether their workers are appropriately aged for the content discussed (Segal & Mayadas, 2005). The use of volunteers interpreters in the longer term can be considered exploitive, as the workers are not receiving support for their efforts and are placing themselves at economic risk for the cause; TWB (2017a,e) highlighted that it is vital to ensure that the workforces in communication services are supported socially and economically. When a communication strategy relies exclusively on the workforce being voluntary it

is considered cheap to run for the agencies and organisations in charge of the strategy, but the deployment of the strategy is costed by the volunteers. Furthermore, another element of exploitation can occur between the interpreters (paid or voluntary) and those receiving support or aid, where the interpreters can exploit the information chain (Carter, 2008). Exploitive interpreters can use their position in the information chain to change the status's of those in vulnerable positions, particularly those receiving humanitarian aid and support. The rumours of such practise have resulted in the questioning of veracity across the migrant crisis in southern Europe (TWB, 2017a), with sentiment of *'get it in writing'* to prove to the next agency both A) what information the migrants are working from and B) where that information was obtained from. Whilst most interpreters, both voluntary and paid, act with integrity towards supporting individuals in need, there are elements within the workforces that do use the situation for personal, social, or economic gains (Maestri & Monforte, 2020). Language policy planners must be aware that the communities are weary of information when they are impacted by emergencies or disasters, due to the displacement and a lack of stability within the situation both physically and sociologically.

A final weakness to consider is that of using children as interpreters and whether they are suitable and appropriate for providing services, specifically whether the age and maturity of the interpreter is sufficient to credibly understand the discourses where potentially graphic or mature topics are discussed (Federici, 2016). It is important when documenting abuses or rights violations that the interpreter understands the technical aspects involved with the types of abuse, particularly when

it is of a sexual nature; and when using a child interpreter, they are unlikely to understand the nuances involved with discussing the topic, such as location of abuse or the injuries endured because of the abuse (Cadwell, 2021). A lack of expertise in the topic can result in miscommunications, both as to the type as well as the style of the abuses, for if the interpreter cannot provide recognisable or specific terms to describe, in the required depth, the abuse itself may not be correctly processed or recorded; as the recorder will have to either downgrade the abuse, to what was clearly identifiable from the interpreter's message or the recorded can fill-in-the-blanks and make assumptions in the recording of the abuses. Both situations result in the incorrect recording of abuses, which can reduce the validity of the recording process whilst also providing opportunities for victims to be mis-treated in long term care plans (Williams et al., 2002). Furthermore, the age-related discussions and suitability are a valid issue that occurs within the active field of refugee management, and as such, is a factor that must be accounted for in wider language and communication planning. There are reports (Edwards, 1998; Carter, 2008; Hart, 2009; Rajan, 2019) of children being recruited as interpreters for conversations about sexual violence and abuses, in some cases involving intimate family members. The discussion of violence is standard practise in refugee processing, as the aid workers ask pertinent questions to identify whether the refugee needs to be placed in a safer region of the aid camp, requires additional mental or physical care and support as well as to keep a log of the wider situation by reporting the crimes. The issue here, is not the discussion of the violence, rather it is on the whether it is acceptable for an

infant to function as the communication go-between in discourses where victims of sexual violence are asked to describe, in detail, the violence that occurred to them.

Overall, the functionality of interpreters is regular; with communication facilitation being the primary role (Federici, 2016). However, the deployment and utilisation of the interpreters services, workforces and contexts is complex and requires attention when planning language and communication strategies for situations of relocation, displacement, and emergencies. The preparatory nature of the communication strategies is another avenue that warrants investigating, as the interpreters are the adapting human language resource that can be deployed in an emergency both on an ad-hoc and prepared basis. When the language strategies are pre-planned for the emergency event, then the alternative source of language resources is the pre-made messaging that can be produced, and as such, the following section will explore the mechanisms for pre-made language resources as well as the technical aspects that have been identified as vital to successful messaging, from a language perspective itself.

Preparing for emergency and disaster events is a critical element to successful management of the people and the mitigation of risk for the impacted population; a statement that includes the language support and provisioning. To facilitate this approach, communication strategies and language resources can be prepared in advance, for the contingency of an emergency or disaster event. When preparing emergency response strategies, it is important for policy makers to consider Sandman (1989; 1993)'s Risk model for communication.

2.3.1.7 Realisation of emergencies and the social responses

In a disaster situation, communication provided is designed to reduce the risk of harm to the overall population and the individual themselves (Kuligoski & Omori, 2014). For this to work the population needs to understand the severity of the situation occurring and have a sense of agreement with the authority that has stated the level of severity. Sandman's (1989; 1993) model of risk management and believability of risk when it occurs is focused on identifying why a group in a community would reject any management or evacuation orders in an emergency event. The Sandman (1989; 1993) model of risk proposed that in order for a population to understand the level of risk from a disaster event, they must understand both the extent to which the hazard is occurring or will occur, and the level of community outrage must be high enough motivate a change in the population. This was specifically designed in mind to explore general compliance of disaster management orders in active situations. Sandman (1989; 1993) theorised that even if the severity of the situation itself is high, if the population is not onside with the jurisdiction involved in supporting them, then they are not likely to follow any orders or comply with any regulations.

Sandman (1989; 1993) conceptualised risk as the overall level of danger understood in a population, specifically related to the extent to which the community will understand and respond to any disaster event or emergency. Furthermore, a hazard is conceptualised as an event that can create a danger to the surrounding local environment; a hazard is one mechanism that result in the cascading effect for a disaster event (ibid). Additionally, outrage is conceptualised as being the level of unhappiness in the population towards either the actor or the event that is leading to

the increased risk in the area (Sandman, 1989; 1993). At the upper end of hazard are events which will create widespread injury to either the population or to the infrastructure, such as explosions or wildfires whereas the upper end of outrage is when the communities hold great amount of discontent the point where they will fall into action without requiring assistance (Sandman, 1989; 1993; 2012). Inversely, if the hazard is perceived as low, which can happen in events which are out of the perceived area, then the risk-to-life can be increased for the population. For illustration, consider a remote rural villager who cannot understand that a fire currently 20 km southwards will reach them in a short period of time if the prevailing wind changes direction. The villagers could reject the notion that the hazard itself is relevant to them and in doing so they would reduce the perception of risk that they have towards themselves and their local community from the specific disaster event.

Furthermore, if outrage is low, then the greatest response given is apathy where the population is not motivated to perform any actions or change their risk of harm from the disaster event or emergency (Sandman, 1989; 1993; 2012). It is important to point out that both hazard perception and outrage can be low, in which case the risk itself is perceived by the populations will be minor, which can contrast directly with the formal assessment of those looking at the situation from either a bigger picture or more objectively. Examples of situations where the hazard has been conceived as low by the population, whereas the assessment risk is high, are in negligence in industrial regions, whereby corporations in charge ignore safety rules as they perceive them to be unnecessary and cumbersome, and when it goes wrong it will result in a disaster, such as the negligence in storing ammonium in Beirut 2010's,

which resulted in the Beirut explosion of 2020 (Al-Hajj et al., 2021). Alternatively, hazard and outrage can both be high at the same time, which then creates a situation where the risk and the perception of risk are greater than the sum of both the hazard and outrage combined. Popular examples of when the hazard and outrage are high included riots, where individuals are compelled to rush to their properties or businesses in an attempt to protect them from what they believe is a significant risk (Sandman, 2012). Riots are a clear example, as the hazard is the people acting in an angry fashion and the outrage is the direct cause for the hazard itself in doing so it is a clear escalation from a minor outrage to a major hazard which subsequently creates a significant risk perception and assessment.

For a disaster management to work efficiently, Sandman's model (1989; 1993) supports the arguments that both the hazard and outrage need to be high. The methods to improve the perception of both the hazard and outrage are different for each factor (Kreps & Bosworth, 2007). Hazard requires the population to understand the properties of the event itself such as the likelihood that injury will occur or fatalities may result, it also requires the population to be educated the level of understanding why the event is an unacceptable risk to the general day-to-day lives and activities of the affected area, in that they need to understand specifically why they need to change their location or take action in response to the event (Sandman, 2012). Outrage on the other hand, requires the community members to feel a negative sense of discontent at a situation that is occurring, and thus is difficult to produce through education alone (ibid). To improve the community outrage, action needs be taken to induce a sense of unhappiness or rejection of a status quo of

doing nothing, which can be done through media, whether conventional television and radio or through social media, where the commentators could provide a sense of urgency to then use as a motivator for outrage (Scanlon, 2007). A sense of urgency selected characteristic for successful evacuation orders and disaster management techniques, this is because when a group of people understand that time is limited and that their actions must be done with haste, the compliance rate higher (Sandman, 2012); this is particularly the case when the situation is active or rapidly developing, where the population reacts to sudden changes in either behaviour or location in response to the situation (ibid).

In summary, to improve the perception of hazard in a population, they need to be educated enough to understand the problem itself and the assessed outcomes that can occur more likely than not, whereas outrage and the perception of outrage, needs to be encouraged in a population through social motivation and production of a sense of urgency across multiple communication lines. Both factors require the population to believe that the information they are receiving in relation to both the hazard and the urgency is of a high veracity and from a source that is trying to manage the situation rather than escalate it.

2.3.1.8 Guidelines for emergency communications

There is no universal set of guidelines as to which language resources to produce in advance, nor is there unified approaches internationally as to the language support offered to disaster planning (Lachlan & Spence, 2009; Cadwell, 2021). However, there are a few similarities between nation states for in-action language policies and communication strategies that have been developed, and in some cases tested.

Multi-national examples include premade audio instruction messages (Arai, 2013a), integrated digital telephonic alert systems (Carter, 2008); traditional media (radio and TV) news dialogue guidance (Scanlon, 2007), digital transition services (Britton, 2007) and applications (Edington, 2021) & pre-written multilingual general information guidance (Frederici, 2016). To discuss and review each of the communication strategies listed above would exceed the scope of this thesis, as the deployment of premade resources was not a significant element identified from the contexts described by TWB (2017a).

A cross-over element that impacts both pre-made and ad-hoc communication is the message construction and style. Across the literature, and historical examples, there were a few over-arching themes and techniques that were identified as being beneficial to effective emergency communication. As such, the following section will explore and outline the four core areas that were identified from the literature, which are: Contents, Associations, Length & Delivery (CALD). This overview will function to provide an understanding of the requirements for emergency communication, the techniques which are accepted as beneficial for information to a population en-masse whilst also supporting a criterion for the construction and formation of any replication of emergency communication for experimental testing or analysis.

2.3.1.8.1 Contents

The contents refer to the information of a message itself, as in the specific targeted knowledge needs to be delivered across the environment to all persons (CDC, 2014). There were five elements that were identified as generalised criteria for effective emergency communication: veracity, urgency, jargon, essentialism, and semantic

ambiguity. The guidance for each element varies, and as such, the following section will briefly outline the features that have been identified as impactful to the effectivity of communication, from both negatively and positively perspectives.

2.3.1.8.2 Veracity

Regardless of the situation, the messaging needs to be truthful and correct at the time of delivery (Kuligowski & Omori, 2014). The veracity of a message cannot be considered a variable factor, i.e., negotiable, as the consequences are serious and at the worst, life-threatening. When people are presented with truthful information consistently in the disaster management and emergency communication, there would be little logical reason for the populations to distrust the information source or the future communication from said source (Sandman, 2012). However, if the truthfulness is scattered and misinformation (due to factual inaccuracies) is relatively common, the result is that the impacted population becomes distrusting, sceptical, and doubtful towards information received from official sources. The worst-case scenarios include a population detaching from the official operations in a disaster event, whereby they reject all information and actions from the officials because of the lies and miss-information that they have heard previously. Widespread rejection of an organisation or government can be explained using Coombs (2007) Situations Crisis Communication Theory (SCCT), which details that the reaction of a community is connected to the pre-existing reputation as equally as it is to the active response. For instance, Libya in the Arab Spring, there was societal distrust towards the government officials, due to corruption and authoritarianism (Zuber & Moussa, 2018); when protests escalated into riots (and therefore an emergency), the response of the

government was to suppress opposition (Aghayev, 2013). The SCCT model when applied to Libya, would indicate that the combination of a starting negative reputation of the government and a negative response to an emergency, would result greater rejection of the authority. In reality, the situation escalated into a civil war and a breakdown in society (Aghayev, 2013). The relationship between distrust and compliance with emergency responses is well-established, and a lack of veracity in messaging is a factoring part in the equilibrium (Calhoun, 2008).

2.3.1.8.3 Urgency

In line with Sandman's (1989; 1993) model of risk communication, effective messaging requires the urgency of the situation to be clearly identifiable throughout the messaging interval. To achieve this, the Centre for Disease Control and Prevention (2002, 2014) recommends using words associated with severity to highlight the risk to life in the environment, terminology includes: danger, risk, and harm. The use of terms that are connected directly with severe and dangerous contexts can allow for the situation to be registered by the population. Projecting a sense of urgency can also be developed in tandem with the delivery system, as word choice itself is not enough to convince populations of risk as a standalone factor (Hellier et al., 2002; Suied et al., 2008).

2.3.1.8.4 Jargon

Jargon, or technical vocabulary, is often recorded across communication sciences as a way to reduce the effectiveness of discourse retrieval when discussing matters to those outside of the field (Bean et al., 2015). In disaster management this issue is

elevated in ways not often considered in the general literature. During peacetime, conveying messages to the lay person, i.e., without technical vocabulary being used in a way to confuse the listener/reader, en masse is difficult to achieve (Lachlan & Spence, 2009). In the pressured of emergencies, the impact of jargon can result in a risk of harm for those who cannot understand (Sandman, 2012), as the information transmitted during the event is likely to be aimed at reducing risk or supporting the alleviation of the situation itself. An additional factor to consider in emergencies is the outrage effect (Sandman, 1989; 1993; Ng & Hamby 1997), the extent to which a population can ensure instability and community trauma due to emergencies before reacting in a negative manner towards those around or within the situation (Sandman, 1989). The use of jargon can disenfranchise the population from connecting with or trusting aid workers or officials, and as a result the populations may take incorrect actions that increase the risk to harm or undermine the relief efforts overall (Fokaefs & Sapountzaki, 2021). Furthermore, if the population are receptive to the message, but fail to understand the information correctly due to the language barrier from jargon, then there is less guarantee that the appropriate response to the information will be taken; this is particularly important with relocation or evacuation orders, where the correct responses are the following of the directions to relocate to a safer or more stable environment. An example of a communication barrier cause by jargon explored using human-conflict scenario, an urban residential area with an incoming missile attack. The messaging for evacuation and order to find bomb-proof shelters could either include the technical terminology for the missile attack itself (example message: incoming ballistic aerial missile strike) or a simplified

description of a non-descript missile attack (incoming aerial attack); in which case the

latter would be favourable in informing the general population as it is unexpected for

the population to understand the military term ballistic. In this example, the

population would be able to register the risk of the attack from both messages, as the

attack is labelled as such, but the chance of misunderstanding is reduced with the

omission of jargon.

2.3.1.8.5 Semantic complexity

Terminology used in emergency communication should be selected using

simplification as a primary criterion, that is to say, that policy makers should favour

simplicity over specificity in messaging. The rationale for this is like the avoidance of

jargon; the message needs to be understood as quickly as possible with little chance

for error occurring from the words themselves (Sutton et al., 2014). To achieve this,

messaging should be designed with as little information as possible to relay the

general information requiring transmission, this can be realised through message

depth truncation and widespread omission of semantically relevant words (Lachlan &

Spence, 2009; CDC, 2014). Lexical items that relate to abstract concepts or ideas

are best avoided in producing emergency communications, as the detailed added by

the words is immaterial towards the overarching aiming of emergency messaging;

which is clarity with speed; example concepts to omit include: metaphors, colours

and (rhetorical) questions (Lindell & Perry, 2003). As a generalised rule of practise,

adjectives and adverbs are avoided where possible, as they both function to improve

understand and depth of core pieces of information (Sandman, 2012); in an

emergency message the core information should stand out independently without

additions. An example is 'residents need to make their way rapidly to the tall circular shelter'; which contains additional detail beyond the essential bare minimum. For the core information is [residents], [go to] and [shelter], as such a truncated alternative message could be 'residents go to shelter', which reduced the length and complexity of the message, therefore making it clearer and more understandable en masse. The use of additional verbs [need], adverbs [rapidly] and adjectives [tall circular] to provide message depth is advantageous, and may not reduce the understanding of some, but it is not guaranteed to be understood in total by all involved. A clarification here, is that this example assumes that both the shelter location and the threat is known and accepted by the population; if the situation was less simple, and the population were less informed, then the addition of more detail in the message would be warranted, but only added to until the bare minimum was achieved. Overall guidance supports limiting both depth and number of additional details provided in all emergency communication, as each level of detail can result in communication difficulties or miscommunications.

2.3.1.8.6 Associations

When individuals hear a message, there can be underlying associations identified or formed in the mind, which can both improve the understanding and acceptance of the information, or alternatively, support the rejection of the message itself (Coombs, 2007). In emergency communication, when exposed to a message an individual may forge connections between pre-existing pragmatic and cultural customs and the message content itself (ibid). When developing and discussing emergency communication, including interpretation, policy makers should understand the

nuances that support or discourage the reception of the messaging, particularly when the messaging contains vital information aimed at reducing the risk of harm. Within this theme, there were four elements identified as being impactful in emergency communication; they were: humour, cultural sensitivity, authority, and positivity.

2.3.1.8.7 Humour

Humour when mis-interpreted or received as an insult, can cause communities to distrust officials (Doğan, 2007), and any person associated with said officials, this is due to humour being an abstract concept, which is subjective to the listener (Bassnett, 2013). To be humorous, it requires an individual to use socio-cultural norms or customers as a backdrop for comedic purposes with the intention to produce laughter or stress-relief. Whilst stress-relief may seem an admirable goal to aim for in emergency communication, given that the individuals involved are stressed, the act of using humour is warned against (Federici, 2016). The abstractness of humour can result in misunderstandings which reduce the effectiveness of messaging and compliance thereof (O'Brien, 2016). Furthermore, the subjectivity of humour can result in negative response, in peace time it is established that comic routines can be poorly received by an audience, such as member feeling offended by the context or implications of the jokes during the routine (Butler & Stoyanova Russell, 2018). In emergency situations, the risk of offending a person from a poor joke can have serious consequences; for instance, if in evacuation orders, the evacuees feel offended by an off-hand joke from an aid worker, the response could be to reject the evacuation orders altogether, or to resist further instruction from either the impromptu comedian or, worse, their organisation

(Coombs, 2007; O'Brien, 2016). In which case, the act of humour has resulted in social rejection, offence and a break-down in community communication and trust. Best practise in emergency communications is to avoid any, and all, instances of humour, as the risks (social rejection) outweigh the positive effects (minor stress relief).

2.3.1.8.8 Cultural sensitivity

When interacting with a community, it is important to follow the customs and traditions to avoid causing unintended distress or offence (Jacobsen & Landau, 2003). In emergencies the same principle applies, but to an altered degree when compared to peacetime, social norms can be broken by emergency communication, such as interruptions or fast-paced speech (Cadwell, 2021), in a manner that is not considered overarchingly rude or discourteous, whereas the same actions in peacetime are unacceptable (Sandman, 2012). However, to assume that cultural sensitivity is not required in emergency communication, as certain conventions can be ignored, is overapplying the notion that the greater-good is at stake. In reality, the sensitivities afforded to minorities and community groups needs to be sufficient to avoid widespread indignation or offense being felt as offence can lead to the rejection of messaging from aid workers and government agencies. Methods to reduce the impact of any cultural variation or schisms, include avoiding the use of nationality labels (Staub, 2009), colloquial labels (Nassenstein, 2019) and referring to anyone as an enemy (Staub, 2009); these examples highlight the sociological framework that exists in communities, and how describing a group with an undesirable term can alienate those involved, particularly if the community considers the term a slur or

insult in their local community (Segal & Mayadas, 2005). Balancing cultural sensitivities, as a task, gets steadily more difficult with the inclusion of more social groups, which can be religious, linguistic, sociological or nationality motivated (Lindell & Perry, 2003). It is important to recognise that human disaster events may result in an additional level of difficulty for the communication planners, that of having refugee's or residents who are from the communities that are against one another, in which cases, the sensitivity of language use in communicating with either or both groups must be placed of paramount importance (Federici, 2016). The consequence of offending a single group in this context, can result in the message deliverer as being considered an adversary, and this negative perception could then become associated with the organisation they represented, which can undermine further work between the organisation and the population (Coombs, 2007). Regardless of the situation, sensitivities need to be considered in relation to the socio-cultural norms of the affected population.

2.3.1.8.9 Authority

A lack of authority in a message can be attributive to a rejection of information (Sandman, 2012), as the listener/reader can reject information as being incorrect, incomplete or disreputable based on the level of authority associated with the messaging. Thankfully, the inverse is also possible, with a strong sense or presentation of authority resulting in lesser rejection of information, and a greater sense of trust between the author and the recipient (Coombs, 2017). In an emergency, features that are associated with authoritative weakness can resonate with the population in a negative manner, as it can invite challenges to the power and

authority of the information source in the discourse (Sandman, 1989; 1993); this is particularly important with pre-recorded resources where speech errors and discourse markers are unexpected, as pre-recorded is structured, formal, and rehearsed. An example of weakness features is discourse markers and fillers, which are units of empty meaning produced to fill space or time in the discourse, such as 'um' or an elongated schwa inserted mid-sentence (Sorensen & Sorensen, 2007). Jones & Askew (2016), highlighted that the use of discourse markers in emergency orders reduces the reliability of the instructions given, and the authority of the speaker was open to challenge, or at least perceived as challengeable, within the population.

2.3.1.8.10 Positivity

Negativity and negation are features discouraged in disaster communication, as both can impact the response to the messaging (CDC, 2012). If a message contains negation, in a grammatical sense, then there is an additional level of processing that would not exist if the message was singularly positive (Maheshwari & Rajan, 2016); and given the rule of emergency communication is to reduce the level of depth of detail until only the essentials remain in the message, negation can thus be cut. There is also a greater chance of miscommunication when negation is used in messaging, as the relationship between the positive element and the negative parts needs to be retained in the listeners mind (Maheshwari & Rajan, 2016); for illustration, consider the following example: Walk down the left-hand lane, not the right-hand lane. If the negation is errored or misheard, then the listener would contain two pieces of information of equal comparison, as both the left and the right side of

the lane could be the appropriate direction; and as such the chance of selecting the correct option is half. Whereas an alternative sentence using only positive elements, without any negation, would be walk down the left-hand lane, which contains less information, yet the objective outcome is the same. Negativity is also discouraged as the use of negative terms, phrase or topics can cause further psychological stress on the individuals, particularly if the individuals are displaced or have endured significant disaster events (Pettit et al., 2015). An example of positive techniques in disaster communication is the avoidance of any negative phrasing, terminology, or topic in sum; so instead of referring to the community's losses; which can be humanitarian, economic or social, the topic focus should be on building a-new, with the overall focus on positive elements that are rooted in the present or future. Grammatically, this can be achieved by using only the present tense in active emergency communication, to denote the need of action in the present, alongside the use of the present and future tenses during the processing and management post emergency (CDC, 2014). A relatively simple instruction for communicators, which has historically been credited with supporting the compliance with evacuation orders and the rebuilding of communities post-disaster (Arai, 2013a; Kuligowski & Omori, 2014).

2.3.1.8.11 Length

Communication in emergencies is time-limited, in that the situations are unstable and the responses to the events occurring need to be closer to instantaneous where possible (CDC, 2014). Alongside the less-is-more approach, there were two other elements that were identified from the research as being guideline best practise for

emergency communication; they were: production length and syntactic complexity. The following section will overview the two elements for greater depth and clarity.

2.3.1.8.12 Production length

The length of messaging is an accepted impactful factor for disaster and emergency communication (Kuligowski & Omori, 2014), with research focusing on finding the goldilocks range or the perfect temporal-window (Arai, 2013a; CDC 2014), where the length of messaging is optimal for information delivery and retrieval. A generalised approach quantified both spoken and written messaging equally, with the length being based on the spoken replication of the message in a single uninterrupted natural speech event (Kuligowski & Omori, 2014). This allows for multi-modal comparisons between messaging systems to be conducted; so, the length of a written message is measured by the production of the message when spoken, as verbatim, as if the message was designed to be spoken originally.

Across the literature there is general agreement that emergency messages be a maximum 30 seconds in length (Arai, 2013b; CDC, 2014; Kuligowski & Omori, 2014), however, beyond that distinction there is dispute and disagreement. When replicating experimental stimulus, to test the understanding and efficiency of messaging, there are disagreements both to length for practical stimulus, as well as the measurement units. There are four units that were identified from the literature, that are used in competition: seconds, morae, syllables, and word count. The generalised de facto measure is word count, as it is used by nation states for developed emergency statements (O'Brien et al., 2018) and orders, however, the three other units highlight

different elements of measure which can be beneficial to briefly consider in an overview.

When seconds are the metric, the literature in this thesis is borrowed from non-disaster explicit sources, due to a lack of applicable studies, as such the literature from Arabic sociolinguistics was explored. Both Koppensteiner & Lenz (2020) & Mirshahidi (2017) support lowering the goldilocks range to between a minimum of 10 and a maximum of 20 seconds. Furthermore, Nejjari et al (2019), proposed the use of speech samples of 10 seconds length; arguing that natural speech sentences were typically around such lengths. Sentences of less than 10 seconds have been identified as being sufficient in length for listeners to identify accent (Derwing & Munro, 1997) and native language (Munro & Derwing, 1999), as such Nejari et al.'s (2018) argument is supported, additionally, there is little evidence that messaging of between 10 and 20 seconds is lesser effective in emergency communication than messages at <10 seconds; highlighting the disagreements across the field.

Morae, or syllable weight has been used in disaster communication research directly as a control measure for stimulus (Hodoshima, 2019). The studies which were found to use morae as a measurement are directly related to emergency communication, as the focuses were investigating the responses to verbal disaster warnings through public announcement systems. Hodoshima (2019) & Ofuji & Ogasawara (2018) used morae as a measure for controlling the length of verbal stimulus, alongside word count. Hodoshima (2019) investigated the effect of words limited to 4 morae inside a controlled carrier phrase, highlighting how the measure can be effective at controlling for specific conditions; whereas Ofugi & Ogasawara (2018) developed stimulus which

was recorded in morae, rather than produced with a specific length as a criterion. The two studies highlight how morae can be used as a measure, but it is worth noting how the measurement has not been adopted beyond the two applications shown. In fact, both morae studies make use of word count as a measurement for sentence length, alongside morae, which supports the notion that word count is the de facto measure for sentence length in both general communication studies and emergencies. When Hodoshima (2019) and Ofuji & Ogasawara (2018) are considered in terms of seconds, the longest stimulus sentences used were 5.6 seconds, adding to the evidence that messaging, and testing thereof, can be efficient with less than 10 seconds of sentence length. As such, morae is a usable, but alternative for measuring sentences for emergency communication.

When considering emergency communication length by syllable, the Centre for Disease Control and Prevention (2002, 2014) supports reducing the syllable length to a maximum of one hundred per full statement. This was motived by average American syllable speech rate of six syllables per second (ibid); which would support the maximum limit for an emergency statement at sixteen seconds long on average (CDC, 2014); however, further guidance advised that slower than average speech should be used (Fokaefs & Sapountzaki, 2021), which would reduce the speech rate. From this, there is an estimated goldilocks range for communication ranges between eighteen and twenty-two seconds in length (Cadwell, 2021). Within the research advocating for syllable length as a primary measure for emergency communication, there is a degree of arbitrary application of the measurement (Arai, 2013a). However, the CDC (2014) approach contains a defendable position, with the syllable length

explicitly linked with the syllable speech rate of the local population; however, there-in lies an issue for language policy planners; that of linguistic variation as syllable speech rate is a factor regulated by both the language of use (Ghazali et al., 2002), but also the speaking community itself (O'Brien, 2016; Federici, 2016); and thus to assume that all speakers from the same region will speak at the same rate is dubious, likewise is it equally ill-advised to assume that all speakers of the same language, regardless of origin or location, will speak on average similar.

Similarly, word count has been used as regular measurement for emergency communication through the literature; however, there are major disagreements as to the criteria for application and the maximum threshold for a stable and effective sentence structure. The only agreement that was identified was that lower word counts were advantageous and to be strived for (Hellier et al., 2002; Lachlan & Spence, 2009; CDC, 2014; Cadwell, 2021). The application of word count is weakened with a lack of agreement as to how to count the individual lengths of words themselves, in fact there was little guidance or research that considered whether longer words should be treated equal to short length counterparts. Furthermore, the application of word limits is arbitrarily deployed, with Bean et al. (2015) supporting a maximum of fifty words, the CDC (2014) now advocating for thirty yet previously arguing for one hundred (CDC, 2002); thus, not providing support for a single threshold as yet. The closest to agreement of word length was found with recent literature on SMS alert systems (Bean et al., 2015), with the word limit being converted into character limits, as the SMS messages themselves are limited in file size for transmission. The SMS character limit is less of an agreed-as-effective limit,

for the goldilocks range, and more of an operational maximum from the specific

message system conditions, however, it can still inform policy makers, as the

deployment of a standardised messaging structure for emergencies should be

applicable multi-model to cater for the developing technology level across the world.

It is important to recognise that focusing on SMS messaging singularly is to

disadvantage the lesser-developed populations (Britton, 2007), which are currently

the most at risk (Vij, 2022), therefore a balance of modern and traditional mediums

should be considered when developing language strategies and communication for

emergencies.

2.3.1.8.13 Clarity

Across the literature, there was agreement that simple sentences are advantageous

in emergency communication. The defence for using simple complete clauses is

motived from comparison studies that found that, when compared to complex and

compound sentences, simple sentences benefit from requiring less working memory

to process, a lower natural error rate with understanding the original message and

correcting when the speech is impaired or damaged (Lindell, 2015), and the being

easier to repeat en masse in emergencies (CDC, 2014). Additionally, the literature

also presented agreement for use of a single sentence (Sorensen & Sorensen, 2007;

Sandman, 2012), where possible, to distribute all information, as the primary

structure for messaging in emergencies, particularly for active situations.

Furthermore, the CDC (2014) supports the avoiding of changing messaging within

quick succession, as each announcement of information can override the precious

message, especially in active situations, which can result in the impacted population

following the latest instructions, rather than considering both and then assessing their options for response. The distribution of information in a single sentence requires the grammatical features to be minimal, as well as the syntactic structure to remain simple; however difficult to produce, the field work result, from historical situations, do support the limitations (Staub, 2009; CDC, 2014; Manuel, 2014; Maheshwari & Rajan, 2016).

2.3.1.8.14 Delivery

The manner in which a message is delivered is an element of communication that is impactful regardless of the setting, however in emergency communications it is vital that the style and approach of a message is not negatively received due to the delivery. Within the theme of delivering a message, there were four elements that were identified as being impactful for facilitating effective communication; they were: imperative, repeatability, gender and urgency. The following section will outline the rationale for each element listed, with examples provided throughout.

2.3.1.8.15 Imperative

The most efficient structures identified for delivery are those which use imperative forms (Inoue, 2012). When a message is received in an imperative form, the urgency of the message is clear from the offset (CDC, 2012), as the form itself directs the listener to act in accordance with the message's directions or instructions. Evacuation guidelines, across the literature, present imperativeness as a vital component for successful compliance with messaging (Kuligowski & Omori, 2014; Cadwell, 2021), regardless of the population demographic. Imperative forms are

used widespread in emergency response systems for both natural and human emergencies across the world; as such imperative forms are a stable, reliable and effective criteria from field work evidence.

2.3.1.8.16 Repeatability

Disaster and emergency communication require the repetition of messaging to ensure effective understanding of the information transmitted (CDC, 2014; Kuligowski & Omori, 2014; Maheshwari & Rajan, 2016). The repeatability of messaging, without deviation, is vital for both pre-recorded and ad-hoc announcement systems and alerts, to facilitate this regularity, the messaging requires a standard structure and form (CDC, 2014). In fact, Arai (2013a) highlighted how a unified structure for emergencies improve cross-linguistic compliance with the messaging, as well as the response to risks and hazards, in a similar way to universal labelling of dangerous industrial chemicals and condition. The Japanese government's disaster guidance for tannoy systems (Fukami & Hisamoto, 2010) includes a premade structure for announcers to base their speech on; likewise, in the USA, the tornado warning dialogue guidance for TV networks (National Research Council, 1991) contain exemplar scripts that presenters should memorise for when declaring emergency alerts for sudden weather events. The replication of a single message has been found to improve the receptiveness of a message, as the sense of urgency is increased with frequency of repetitions (Sandman, 1989; 1993) as is the understanding of the messaging (Hellier et al., 2002) and the compliance with the specific information in being transmitted (Kuligoski & Gwynne, 2010).

2.3.1.8.17 Gender

Cross discipline disaster communication and intelligibility research can inform the gender of message delivery in emergency communications. In intelligibility research it is established that female voices are more intelligible and understandable when heard (Kwon, 2010); with lower error scores for understanding noticed across a range of studies and conditions. Although the root cause for this phenomenon is unknown, the resulting improvement in comprehension and intelligibility can be used to improve the delivery of emergency messaging. Furthermore, recent research in disaster communication has used female voices for message delivery, crediting the stability of evidence supporting female voices as more intelligible (Hodoshima, 2018; Ofugi & Ogsawara, 2018; Smith et al, 2019)

2.3.1.8.18 Urgency

In line with Sandman's theory (1989; 1993; 2012), the urgency of a message and situation are vital elements of effective communication between official agencies and the impacted populations. Sandman (ibid) produced this theory based on his own field work experiences, and research since as widely agreed with his approach. Since Sandman proposed his original theory, research has identified that urgency is comprised of multiple elements (Suied et al., 2008), many of which are related to the spoken presentation and delivery of a message; specifically, the production speed, prosodic pace, and the general fluency of the speaker.

Earlier research into speed identified that when listeners heard messaging, if the delivery speed is perceived as faster than naturally occurring (Carter, 2008), the

comprehension of the message is increased, alongside an increase in the sense of urgency when scored; a trend which continues in modern research result too (Ofugi & Ogsawara, 2018). Likewise, a quickened pace of the message, when spoken, equally improved the sense of urgency felt by the listener (Lam & Tjaden, 2013). The extent to which a message should be quickened is still debated, with research agreeing only-so-far as that faster than average, but shorter than unintelligible is as a *goldilocks range* thus far (CDC, 2014). Furthermore, fluency of a speaker is also related to the sense of urgency from an emergency message, specifically the level of disfluency. Wogalter et al. (2002) identified that the higher the production error rate of a speaker, the weaker the sense of urgency in the listeners; this relationship is considered a factor of social perception (Li, 2004), as formal contexts typically require, or at least are expected to require, a greater level of precision as opposed to informal contexts where impression is not unexpected or uncommon. The agreement across the literature is that messages should be delivered with a faster-than-average discourse speed with effort made to reduce the speech error rate to null; however, with pre-recorded the goal should be easier to reach, as the message can be repeatedly recorded and crafted to ensure there are no errors, whereas the best of effort approach is advised for ad-hoc live messaging.

2.3.1.8.19 Conclusion

When preparing disaster communication for widespread usage, it is vital that the message itself contain the least amount of information as possible to achieve the target goal. The information provided should be direct and provide efficient evidence for the population to comply with the message, which is particularly vital in

evacuation orders, without questioning whether the statement itself is direct. There is little room for abstract notions which reduce the directivity or bluntness of the message if the aim is to improve compliance and following of the information. It is also important that the message itself be clear in a simple manner, by this, the information given should be presented in the least complicated format possible, to limit the depth of information can be transmitted. Though it may seem counterintuitive to advocates towards less information being presented to those at risk, the counter point is that using more information in fact reduces effectivity of evacuation orders and Information Systems. Therefore, simple sentences and restricted lengths, such as avoiding messages which are perceived as excessive or long, is an established essential for productive and effective communication in disaster emergencies.

This subchapter has explored the mechanisms for delivering communication in an emergency, as well as the approaches and techniques for both crafting and preparing high-quality and impactful messaging. However, this exploration has not provided insight into why the language situation in the southern Europe migrant crisis is as dire as reported. This sub-chapter has highlighted how the situation itself is not a result directly from the communication strategies being used in the field; in fact, the issues of a language barrier are not found at the communication strategy level of planning, rather the responsibility may in fact be found higher in the top-down chain. As such, the following section will explore and investigate the status quo of current language policy, with focus on the processes for making decisions, language crafting and policy adaption.

2.3.2 Language planning and policy for international situations

This subchapter will investigate the desired characteristics and aims for disaster language policies. The initial sections will outline the general guidelines for identifying and developing a disaster language policy. Following this, the core requirements for disaster policies will be presented, with assessment as to which language-based elements are vital for a language policy in a disaster response framework. After which, there will be a review of the language selection process for international language policies, to identify which approaches can be applied for developing disaster language policies. To achieve this review, an exploration of four current language policies will be presented, alongside, the established issues and criticisms levelled against each policy in turn. The aim of this subchapter is to present and review the process of language selection in language policies, and to explore how current language policies can inform the development of future disaster language policies.

2.3.2.1 Defining disaster language policies

Within sociolinguistics, there has been little attention devoted to exploring language policy, as a topic, within disaster situations. A disaster language policy is a set of regulations, rules and guidelines organising the use of language within a disaster region. O'Brien at al. (2018) explored language access through the term communication framework, but this focused on the human rights access and the way a language was used for communicating. A major part of general language policy is the selection of which language to include (as discussed on pp.51, with the language

orientation paradigm), this is particularly the case for larger policies, such as the nation-state level with the selection of the official language (Ruiz, 2016). In disaster research, the focus is on the manner of language use (Federici, 2016), rather than which language is used, and herein lies an issue. If language use is regulated, those rules and guidelines are dictating the language itself, which is a language policy, regardless of whether the authority enforcing the rules calls it as such (Kaplan and Baldauf, 1997). Therefore, in this thesis, organised or guided language use in a disaster will be considered foremostly as a language policy.

An example of current disaster communication planning, and how it fits as a top-down language policy, is the Japanese tsunami warning systems placed along the nation's coastlines. The Japanese government, in their efforts to reduce the impact of disasters, implemented public announcement systems along the nations' coastlines (Fukami & Hisamoto, 2010). These systems would provide messaging to the general population warning of incoming tsunamis or earthquakes; disasters that occur with a high frequency in the nation. The decision to implement a language strategy, and the method for implementation was decided by the government, and the public were not involved (see figure 1). Furthermore, the warning messages were developed for delivery via the public announcement system, and again, this was conducted by the governing authorities of Japan (Nakanishi et al., 2014). It was decided that pre-recorded alerts would be produced for the deployment in an emergency (Arai, 2013a), so that the system could activate automatically when an alert is issued (Inoue, 2012). The monetary cost of developing and maintaining the public alert infrastructure was funded through the general taxation of Japan, whereby, the

government distributed the finances without discussion from the public (Nakanishi et al., 2014). The public were informed of the alert announcement system during emergency-response training sessions in schools and community groups on what to do during a disaster event (Arai, 2013b). These training sessions are not uncommon in Japan as the nation is highly prone to experiencing natural disasters (Edgington, 2021), and the distributed information on public announcement system through this mechanism was deemed advantageous and cost-effective (Fokaefs & Sapountzaki, 2021). This disaster communication response highlights that it is in fact a top-down language policy, given that the decisions related to language use were prescribed onto the population, in a way where the general public did not have any input on the mechanisms used or decisions taken for the development, distribution, and maintenance of the policy.
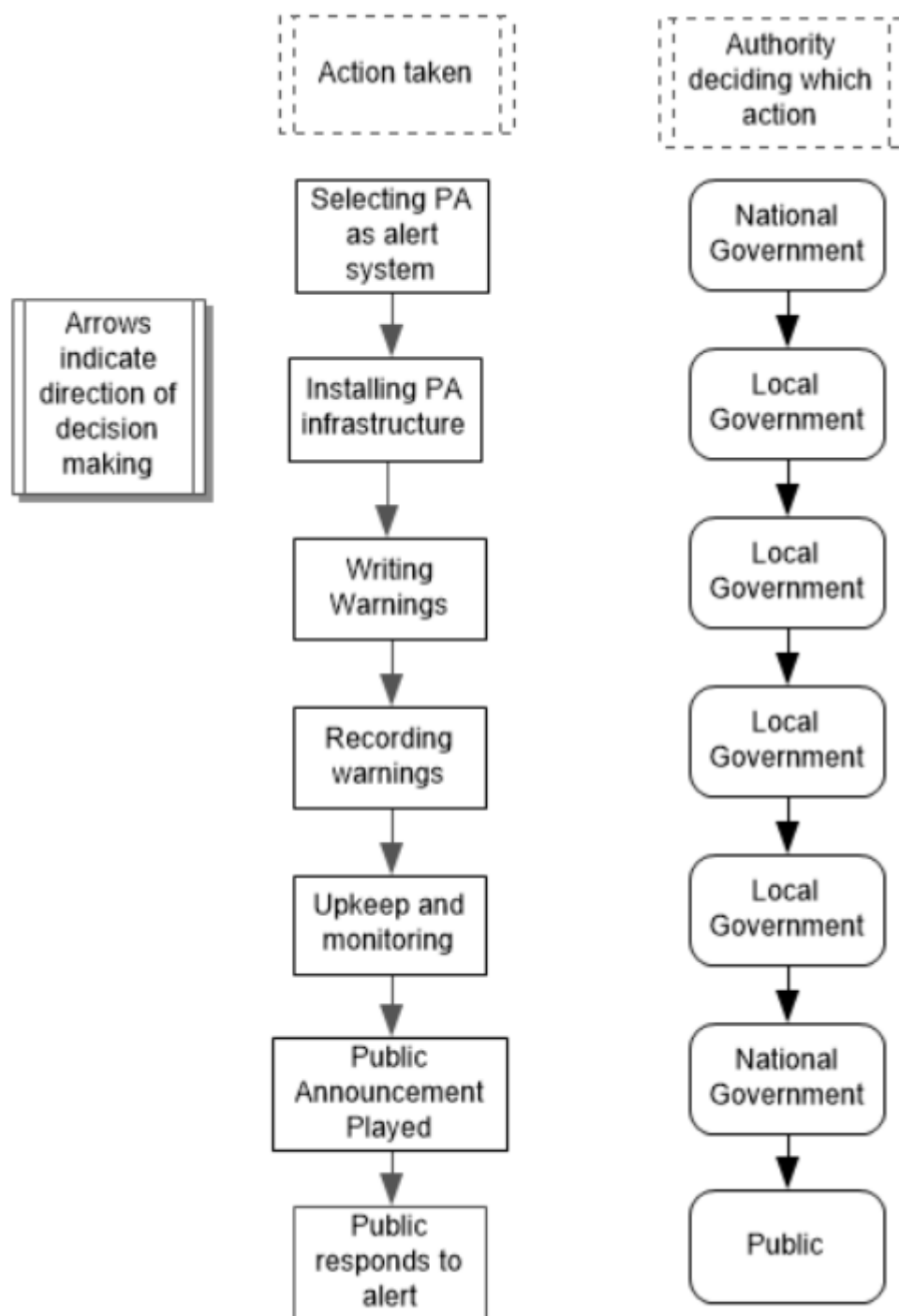
Figure 1: Flow-chart (simplified) of the processes and authorities involved in the deployment of a PA system for disaster alerts in Japan.

2.3.2.2 Core requirements for disaster language policies

Fundamentally, communication strategies designed for use in disaster situations or emergencies are language policies, however there are specific characteristics and requirements for disaster policies in general that need to be accounted for when developing language policies for use in the disaster context (see section 1.2 for first discussion on this topic). Therefore, this is an applied context that is an intersection between disaster management and sociolinguistics. To understand the core requirements of the language policy, it must be reviewed within the context of wider disaster policies and their subsequent minimum requirement. The following section will outline three primary characteristics that were identified from the field of disaster management that can be applied to a language policy to ensure that the policy will be acceptable for use in the field. The three characteristics are that a language policy must be deployable (Britton, 2007), preparable (O'Brien et al., 2018), and socially acceptable (Sandman, 2012; De Varennes, 2021) of which when in combination provide a foundation for language policy be successful in a disaster event. Each of these three requirements will be explained with examples below.

2.3.2.2.1 Deployable

Disaster management as a field is unified in that any action taken by an organisation or a community in a disaster must focus on the reduction of harm and the mitigation of risk-to-life for the population in the affected regions (Carter, 2008). To facilitate this, disaster relief policies must be flexible enough to be deployed into the field at short notice (O'Brien et al., 2018); this reflects the short notice nature of disaster events This is the same for both natural and human-based disasters, and therefore

the field overall advocates that disaster response policies should be completely developed, to ensure that policies can be deployed with short notice (Mayhorn & McLaughlin, 2014f). With language policies, this means that communication techniques or delivery mechanism for the deployment of alert messaging and communication should be finalised and operationally tested prior to use in a disaster response. The majority of language policies are either politically motivated or are stalled by logistic difficulties (Kazeem & Suleiman, 2020). This delay between the selection of a language to use in a situation and the deployment of the policy, so that the language can be used in the context prescribed, cannot be applied in a disaster management field.

A caveat here, is that it is not possible to prepare for every single instance that could happen in a disaster, whether that be the type of disaster, the impact of the disaster or the population affected; however, steps should be taken to make sure that the infrastructure is stable enough to withstand the impact of the disaster (Edgington, 2021). This element of disaster resilience requires policymakers to ensure that there are multiple communication avenues being utilised in a single policy, whereby the infrastructure required for the deployment of the policy is greater than one single technique, such as having printed resources alongside a public announcement system is more suitable than relying on a public announcement system as the only method of distributing information (Britton, 2007).

2.3.2.2.2 Preparable

Resource stockpiles, and resource management, are two key areas required in a successful implementation of a disaster language policy (Cadwell, 2021). This

includes humanitarian policies, where the goods required for supporting refugees in a conflict zone should be available before the delivery of aid becomes untenable in the disaster situation. The development of resources during a disaster is deemed unacceptable in the field of disaster management from an ethical perspective (Ng et al., 1997; Mayhorn & McLaughlin, 2014); predominantly due to rushed resources being of a lower quality in contrast to resources which were developed and tested prior to a disaster. In terms of language, this requirement was recognised by O'Brien et al. (2018), who identified that the quality of interpretation and translation services provided to refugees was directly impacted by the level of preparation and quality control that was undertaken for both resources used and for the interpreters training; whereby, low effort in preparation, or low quality correlated to poor communication in the field. On a wider scale, the languages selected for use in the disaster needs to be those which have the basic capacity to be replicated and therefore translated from (Geiger & Pécoud, 2010; Cadwell, 2021). This in practice means that standardised written languages are more beneficial for use in a disaster situation than oral languages which are not documented. Whilst this may seem contradictory to the notion that languages should not be discriminated against; by which the language community should not be excluded by the nature of the language used; it is important to recognise that in a disaster what needs to be produced is premade resources that can be deployed in a wide application of situations to a variety of language communities. Oral languages cannot be used as variably as written languages, the use of an oral language would limit the mediums of information delivery in comparison to written. For instance, written language can be used for signage, SMS

alert systems and food labelling, without human interaction, therefore saving time for the aid workers and refugees. Whereas, if an oral language was used, all information would have to be given verbally, and this would slow down the disaster response and is prone to increased errors and miscommunication (Lindell, 2015).

At present, the focus across disaster management is on developing strategies and policies that when deployed can reach the maximum number of people for the least number of resources. This creates a practical bias towards favouring majority languages, as the focus is the volume of speakers; thus, the selection of a minority language is unlikely, particularly if the speaker population is also a minority (in perspective to the other languages in the disaster zone), although communicative function for majority is hindered by Ausbaucentrism. It is established in linguistics (via Abstand approaches) and human rights legislation that minority languages are equal to majority languages in terms of status and ability to facilitate communication (Tollefson, 2006; Skutnabb-Kangas, 2009; Wiley & García, 2016), the latter being assuming that an Abstand approach was taken. However, in a disaster, the practicalities take precedence, as the larger majority languages contain greater speaker populations and therefore are considered more effective in providing general language support. However, Ausbaucentrism can limit the impact, as variation and intelligibility issues occur more frequently, whereas an Abstand approach could resolve the issue, with communication placed as the primary criteria, therefore the ability to communicate successfully would be at the forefront. The issue within policies is with whether an Ausbau or Abstand stance is taken when establishing a

policy's priorities (see 2.3.1.4 Abstand and Ausbau for detailed discussion on this topic).

2.3.2.2.3 Socially acceptable

Compliance with messaging and information is a vital aspect of disaster policy as the evacuation of a population from an at-risk zone is a major priority (Kuligowski & Omiri, 2014; Sandman, 2012). When trying to persuade individuals to leave the at-risk areas, disaster management research indicates that effort should be made to alleviate distrust between the evacuee and the disaster worker (ibid). In language policy, it is established that if enacted policies do not conform to the social framework of the population, then that policy is, in itself, at risk of being nullified and ignored by the population targeted (Phillipson, 2006; Smith, 2013; Kruse & Ammon, 2018), so a disaster policy must also conform, in part, to the population. A counterpoint is argued by Sandman (1989; 1993), who argues that trust is not the primary characteristic for compliance in an emergency, rather that is the sense of urgency related to the perception of risk to the individual; however, Coombs (2007) argues that the reputation of the organisation requesting action from an individual is the primary characteristic required for a positive response and therefore compliance.

When considering language use in a disaster, it is important that the linguistic conventions used by disaster workers do not register in the target community as being counter-active or resistant to complying with the social and linguistic norms of their community (Calhoun, 2008). If disaster workers can be perceived as part of the community, then they are more likely to be believed by the community (Bolin, 2007), which in the context of evacuations would mean that the perception of risk to the

individuals would be high, and therefore according to Sandman (1989; 1993) resulting in compliance; and in addition to this the reputation of individual delivering the warning message would equally be perceived as positive and therefore improving compliance (therefore satisfying Coombs, (2007)). It is established in sociolinguistics that prejudicial attitudes towards the speaker, and the message delivered, are influenced by language practices used by the speaker (Abdel-Rahmen, 2016; Abbas et al, 2020). The role of language policy in a disaster, therefore, must account for whether the practices deployed will be perceived as acceptable to the target population, with the mechanism used for dispersing the information being perceived itself (of themselves) as being positive in terms of associations. Furthermore, when language policies used in a disaster do not consider negative attitudes towards a speech community, the results can be unfavourable, particularly when the disaster workers are perceived as being enemies of the community (Staub, 2009), historic examples range from aid workers being assaulted (Jones & Askew, 2016) to those who have been kidnapped and ransomed (Staub, 2009; Jones & Askew, 2016). In general, with language attitude research, the focus is not placed on the negative repercussions of the speaker being severe enough to result in an increased risk-of-harm, however, in disaster research, the focus is on high-risk and unstable environments which can escalate into harm for the aid workers if the disaster-impacted population develop negative attitudes towards them (Sorensen & Sorensen, 2007).

2.3.2.3 Language selection in current international language policies

The current status of language support in refugee management indicates that language policies and strategies are not meeting the requirements for successful facilitation of communication between multiple languages in context (TWB, 2017d).

There are two approaches which can be used to improve the language support in the field, the first is to develop new language policies (Williams, 2011; Spolsky, 2011). There are two criticisms with starting a-new, firstly, is that restarting efforts, with the intention of future disaster events does not address the current issues that are being experiences by the impacted populations at present. The second criticism is that the re-creation of a language policy requires substantial effort for the development and production of language resources (Kudryavtsev et al., 2002; Judge, 2007; Leibowitz, 2015). An alternative approach is to use existing language policies, which have already received effort and resources to developing language infrastructure (Lin & Martin, 2005; d'Almeida & Otco-Grillman, 2013), albeit for different purposes, and adapt said policies for use in a current disaster. This would allow resource requirements to be lower than starting anew, whilst also assessing whether a current language policy is reflective of the population that it is supposed to support or represent. As such, the following section will explore current international language policies, to ascertain if there are differences in language policy production, language selection and the overall approaches taken towards multilingual contexts. The four policies explored are: the United Nation's Language Policy (UNLP), which covers 195 nations; the European Union's Language Policy (EULP), which covers 27 nations;

the Arab Leagues Language Policy, which covers 22 nations; and the African Unions Language Policy, which covers 55 nations.

In a disaster policy, the use of one of the international language policies explored would satisfy the criteria of being preparable, as the resources may be already available (depending on the language), and there exists infrastructure to produce materials in-bulk. Additionally, the languages selected across the four policies are all recognised as internationally valid, by virtue of their inclusion with the respective language policy. This inclusion has granted the socio-political prestige of the languages on the global scale, with their value as a diplomatic language being established (Lin & Martin, 2005; McMenamin & van der Walt, 2018). As such, all the languages used satisfy the criteria of being acceptable, at least in a global perspective. However, the notion that disaster language policies can be developed in such a way to make use of the existing language policies in action, requires policies to be acceptable for use in a disaster situation initially.

At this stage, the general acceptability of using the languages in the international policies is established in a hypothetical sense, as there is little comparison to the language deployment in an active disaster zone. Moving forwards, this thesis will assess whether the current international policies can be used as the baseline for current and future disaster language policies, when the aim is to provide language support to the greatest number of individuals within the disaster zone.

What has not been explored in the research is whether the language policies, and the languages therein, used by international organisations are in fact indicative of communication on a wider scale, as there is a research gap. This gap will be

addressed in this thesis in chapter 3, with the following research question used to guide and investigate suitability of using international language policies in disasters.

(1) Are international language policies suitable for use in emergency disaster language policies?

2.3.2.3.1 United Nations Language Policy

The UNLP started with five languages (English, French, Spanish, Chinese and Russian) in 1945 (UNGA, 1945), with English and French being further classified as the working languages in 1946 (UNGA, 1946). Initially, the UN General Assembly (UNGA) working language policy allowed for only the use of English and French (UNGA, 1946), but over time, this policy expanded to include Spanish (UNGA, 1948), Russian (UNGA, 1968), Chinese (UNGA, 1973) and Arabic (UNGA, 1973). As such, there are six official and working languages of the current UN General Assembly and Security Council. Other languages can be used to address the General Assembly, but only if prior notice is given (Kudryavtsev et al., 2002).

Most of the world's nations are represented in the UN, and the UNLP is the primary policy for dialogue between nation states in the diplomatic discussions (Loos, 2007); additionally, the UNLP is the guiding policy for language resourcing and strategizing for the UN operations and missions. As such, the UNLP is the largest policy both in population catered for and geographical area. The working languages must be used in all discourses, and all official statements, treaties or documentations must be published and made publicly available in the six languages (UNGA, 1973). The UN's

working languages are international languages of diplomacy and well acknowledged for historically facilitating communication between nation-states (Tonkin, 2011). The UN, and the UNLP was designed by the Allied nations in the aftermath of World War 2, with the aim of replacing the League of Nations (ibid) and providing a mechanism for resolving future territorial disputes between rival nation states; to in essence prevent another global war (Bourantonis, 2004).

In the foundation of the UN, first notions of language were in its founding charter, where language, and language access was defined and codified as a human right; Article 55 (UN, 1945) states that there should be 'universal respect for … language'; Article 76 highlights that the UN shall 'encourage respect for … the fundamental freedom of … language...'. Although, the languages of the UN were implicitly decided through Article 111 which outlined that the founding charter must be presented equally in five languages (UNGA, 1945). These five languages were codified as the official languages of the UN a year later (UNGA 1946), clarifying the language position of the organisation. The United Nations approach to language policy highlights the impact that international validation of a language can have on a speaking community, for instance the inclusion of Chinese and Russian was presented by their national media's as endorsement of their culture and identity being superior to their neighbours (Laverack, 2015).

Furthermore, the UNLP started with five languages, and has expanded by a single language since (that of Arabic in 1973 (UNGA, 1973)). The arguments presented to include Arabic as a language for the United Nations (making it a de facto official global language) were to improve the relationship between Arabic speaking nations

and the rest of the global community, as well as boosting and fostering positive

diplomatic statuses (Versteegh, 2014). As such language policy can be used to

create communication bridges and avenues for discussion between cultures and

societies with vastly different standards, statuses, and customs.

As the United Nations was designed to be a diplomatic facilitator, rather than acting

like a government, the policies used within and by the organisation need to provide a

positive development to the key benefactors or partners to continue a peaceful and

cooperative environment between nations (Sokolovska, 2017). It is therefore difficult

to see how the United Nations justify the removal of a language from their

overarching global language policy, as to remove language could be offensive in a

diplomatic sense, and counterproductive to reducing tensions between nation-states

(Johnson, 2013). As such diplomatic language policies infrequently change, and

often when change does occur it is towards inclusion rather than exclusion (Wright,

2016), and so to automatically apply a language policy that is a diplomatic policy, it is

important to remember the limits faced by virtue of its size and scope.

From a humanitarian perspective, the UN aims for equality for all, as well as the

protection of every faith, sex, gender, and belief from discrimination due to these

characteristics (Weiss, 2010). This is also applied to language in the UN charter, as

Article 55 instructs the UN to promote the universal respect for language without

distinction (UN, 1945). Therefore, the general approach of the UN, towards language

use is in favour of including languages regardless of status or size.

2.3.2.3.2 European Union Language Policy

The European Union has a multilingual approach to its language policy, like the UN, and has enshrined the language rights of minority communities within its member states (Kruse & Ammon, 2018). For instance, the EU fully supports the Council of Europes' charter for regional or minority languages (Council of Europe, 1992, 1995, 2010; De Vareenes, 2021), which legislates language support for minority languages found across the Europe, examples include Welsh and Catalan; chartered regional or minority languages also quantify the vitality of languages within the European Union, therefore placing emphasis on the development of existing languages and the retainment of diverse communities within the European Community (Council of Europe, 1992). Furthermore, there is also the framework convention for the protection of national minorities (Council of Europe, 1995); which guarantees the right of a national language to be used within the European Union's working directorate. Whilst the human rights aspects of language are established within the EU, there is also a generalised approach towards encouraging multilingualism within the European commission, which is the executive branch of the EU, as shown in the 2008 communication from the commission to the council highlighting multilingualism as an asset for Europe and is a shared development (Commission of the European Communities, 2008).

In terms of the EULP itself, and the languages selected within the policy, the final list of languages is large (Sokolovska, 2017). There are 24 official (or procedural) languages, and of which there are three working languages which are used for day-to-day communications (EU, 2007), which are English, French and German. Additionally, the nation-states that are members of the EU are encouraged to adopt

their own national language when interacting with the EU formally (EU, 2014), such as delivering speeches in their national language to the European Parliament. Not only is the use of a national language acceptable it is widely encouraged, but there are also no negative sanction mechanisms in the EULP for the use of a national language from a nation-state (EU, 2010; Tomasi, 2017). The EU has multiple objectives for their diversity and integration policies, one of which is the objective to have every EU citizen being able to use two other languages in addition to their native language (EU, 2008). This highlights that the EU encourages trilingualism as a minimum standard of multilingualism, and therefore exposing integral natured multilingualism in the EU.

2.3.2.3.3 Arab League

The Arab league's (AL) language policy is monolingual, in that Arabic is the sole language of use for official dialogues and for negotiations (Versteegh, 2014). The justification of this single language policy is rooted in the shared social identity of all members within the Arab league, unified as Arabs and Arabic speakers. Use of additional languages with the organisation is discouraged (Hachimi, 2013), and in some situations the use of a language other than Arabic can be perceived as offensive or disrespectful to the Arab league overall (Holes, 2011). Use of a single language for a single community is not a new phenomenon, as the one-language one-state one-nation rationale has been deployed throughout history, such as the French unification (Wright, 2016); in this case the argument is one language one people for one geopolitical block (Bassiousney, 2009). What is interesting to note, is that colloquial Arabic's (QA) are acceptable to use in the day-to-day communications

within the Arab league (Albirini, 2011), however the formal written notices and treaties must be in Modern Standard Arabic (MSA) (Versteegh, 2014). The local Arabic varieties are accepted as they are colloquial and are used in the day-to-day running of the nation-states involved in the Arab league (Albirini, 2014). There are many colloquial Arabic's found within the member states and therefore whilst the policy is monolingual, there is a great degree of variation which can be found between the speakers in the organisation (a topic which will be addressed in more depth in a later subchapter).

2.3.2.3.4 African Union

The African Union (AU) is a multilingual organisation made up of 55 member states, and the language policy is codified in article 11 of the constitutive act of the AU (Organization of African Unity, 1963), which is the binding agreement for the organisation. Article 11 states the official languages of the AU are Arabic, English, French, Portuguese, Spanish, Swahili, and in practice, 5 working languages: English, French, Portuguese, Arabic, Spanish (ibid). With these working languages, Swahili is in the process of being added as a working language and should be implemented for use in all communications by the middle of 2023 (AU, 2022); this is a step towards post-colonialism, whereby the languages of Africa are granted equal status to the languages of the colonial empires. However, the AU also identifies the benefits of using European languages not only in their communications but also in their language strategies for education, as the economic benefit of the European languages is greater than the benefits of solely using African-based languages (Juffermans, 2015). The European languages are global lingua francas, as they still

hold dominance in the global aspects of trade, finance, and economics, which are the three areas the AU promotes for the development of member states (Shin & Kubota, 2010; Kaplan, 2018). In addition to the official languages any other African language may be used in the organisation, whether that be in the parliament or in the treaties (Organization of African Unity, 1963). The caveat that any African language can be used, is an insight into the postcolonial era of Africa, whereby populations are encouraging the development of the resources, cultures, and languages native to the continent as opposed to supporting the languages which were imposed by the colonial powers. The AU does not discourage the use of colonial languages, as the right to use these languages is enshrined in the AU's constitution. This highlights the position taken by the AU, as one that states that dominant colonial languages are acceptable for use, and that the right to use them should not be discriminated against, thus the speaking communities of these languages should not be discriminated against. The addition of Swahili as a working language highlights how African language policy is developing and is becoming more inclusive of the population represented by institution, as there are estimated to be over 200 million Swahili speakers which account for 1/6 of the continent's population (Ethnologue, 2022).

The language policies explored highlight how multilingualism is typically encouraged in wide reaching policies and approaches, however this is not necessarily always the case. There are great variations between languages selected for use in the policies, with regional specific languages being supported by the corresponding authorities, such as the EU supporting European-based languages (Ammon, 2006) and the AU

supporting African-based languages (Juffermans, 2015). The crossover languages, i.e., those that appear in multiple policies explored are English, French, Portuguese and Arabic; of which three of the languages were used by the colonial empires historically (Lin & Martin, 2005; Phillipson, 2006; Wright, 2014; Versteegh, 2014); therefore, highlighting how colonialism is still impacting modern language policies. The variation between the four language policies, indicates that there is disunity across international organisations, in terms of the approaches to language selection and justification thereof. The following section will outline criticisms of the four policies, to identify the reasons why variation exists between the policies, beyond the regional language selections.

2.3.2.4 Criticisms of current international language policies

2.3.2.4.1 United Nations Language Policy

There are several criticisms that can be made regarding the United Nations language policy. The first criticism is that the UN favours the status quo, which is a snapshot from post-World War 2 geopolitics, rather than adapting to reflect the modern geopolitical landscape and the changes which have occurred, or are developing, since 1945 (Wright, 2016).

The UN charter (UN, 1945) outlines that a person should be able to receive information from the UN free of discrimination and without being negatively impacted by which language they speak. However, the UNLP suggests that the speaking populations of the six UN languages are officially of greater priority and therefore face lesser discrimination (d'Almeida & Otcu-Grillman, 2013), in terms of language

access and the use of their language, than the minority languages speaking populations (which are not overtly supported). It is acknowledged that to have a language policy including every language in the world is unfeasible (Loos, 2007) as the policy resources will not stretch far enough to include all 7000+ languages globally; to limited to 6 languages whilst also highlighting that language access, as in the ability to communicate with the organisation in your native language, is a fundamental human right, which should not be devalued or worked against by the UN, is problematic when attempting to provide balance for justification.

As a compromise the UN allows for any person to conduct business, provide testimony, or further contribute to the General Assembly or working groups in whatever language they wish, however they must provide advance notice to ensure that the translation framework is in place (Kudryavtsev et al., 2002). Whilst having to give advanced notice is not in itself a form of discrimination, it is a covert barrier restricting ad-hoc conversations to only being accessible in the main six languages. The rationale for why advance notice is required for languages beyond the official six is logistics, as the requirement on translation services and interpreter services is already in high demand within the UN, catering for the current six languages (d'Almeida & Otcu-Grillman, 2013).

Other languages have been suggested for promotion to being an official language of the UN, examples include Portuguese (d'Almeida & Otcu-Grillman, 2013) and Swahili (Williams, 2011), however Arabic is the only language which has been successfully added (UNGA, 1973). The argument for including Portuguese as an official language is born from the sheer volume of speakers of Portuguese, both Brazilian Portuguese

and the Portuguese spoken in Portugal (d'Almeida & Otcu-Grillman, 2013), which is a result of Portuguese colonisation and the retainment of Portuguese as a language within former colonies (Ager, 2001). Contrastively, Swahili has been suggested as a language to include as it is a widely spoken language in several African Nations (Samuelson & Freedman, 2010), and that to add a language native to Africa would signify an acceptance towards identifying Africans as equal to Europeans. This argument may not be directly related to language policy, but it is a reference to the impact of historic social policy, that being colonisation, can have on long-standing approaches taken towards certain sections of society; in this instance nations that were colonised have yet to receive equal standing diplomatically in some circles as the nations which did the colonising (Phillipson, 2006). Resistance to adding Portuguese and Swahili is often described as resisting a change in the status quo (d'Almeida & Otcu-Grillman, 2013), with logistics relating to the resources and expenses required to implement an additional language to the UNLP. As a caveat, there are two primary costs to consider when adding a language to an international policy. The first cost is the staffing, for the language planners, translators, interpreters, and technology engineers (Johnson, 2013). The second cost is the continuous cost of maintaining the translation in terms of services across the UN's missions and diplomatic services (d'Almeida & Otcu-Grillman, 2013).

2.3.2.4.2 European Union Language Policy

With the EU's language policy there are three primary criticisms that were identified; that there is a hierarchy within policy (Kruse & Ammon, 2018), financial cost of running the policy (Ginsburgh & Moreno-Ternero, 2022)), and the great logistic

demands for enacting the policy for day-to-day communication (Ivana, 2022). The

first issue of there being an established hierarchy within the policy, is motivated by

there being an overt EU-wide approach to multilingualism coupled with a minimum

standard of supporting all languages and abolishing linguistic discrimination in all

contexts (ibid). These objectives are at contrast with the social discrimination present

between the languages used within the organisation itself (Kruse & Ammon, 2018).

To use one of the languages found in the policy, which is not stated as the

procedural, requires advance notice to be submitted, detailing who the EU operatives

or bodies that are to be communicated with. Therefore, despite the best intention

towards universal multilingualism the EU itself has produced a bias in the system

towards three languages, which implies that the procedural languages are more

prestigious or useful for communication in comparison to the 21 languages. The

reason to explain why there is an inbuilt social discrimination in the EU language

policy can be due to the criticism that the policy is politically motivated, and that each

nation state has their official language included in the overall policy by default, yet

they may never use their language like they would in their own nation state because

of the need to submit advance notice. This point leads into the final two issues with

the EU's policy, justification for advance notice requirement for the use of one of the

21 official, but not procedural languages, are that it is an expensive activity to have

universal translation between 24 languages of the EU at any one time (Ivana, 2022);

as it would require numerous interpreters the operating in unison. Remember that

when interpreting between languages the interpreter needs to be able to understand

one language and then relay the message in another, and this would be needed

repeatedly for every language contact (Bassnett, 2013). For example, a presentation in one of the 24 languages would require 23 interpreters, with each language matchup requiring adequate language support, as such, coordination of staff, and the cost of the staffing is a justification for the lack of multilingualism in a wider sense in the EU's discourses, parliament and within the commission itself.

2.3.2.4.3 Arab League

The Arab league's language policy is the only policy identified in this study as being against multilingualism, by favouring a single Ausbau language for use in every context. This absolutism is a major criticism of an inter-national language policy and organisation, as multilingualism is the default standard for the global perspective, as there are more bilingual speakers than monolingual speakers (Bardel et al., 2013). There is also the issue of intra-language variation (a common issue of Ausbaucentrism as discussed in *2.3.1.4 Ausbau and Abstand)*, as colloquial Arabic's are acceptable to use in the official discourses despite the recognition by the Arab members that communication and intelligibility between Arabic's can be impaired if the distance between the Arabic's is too significant (Zarka & Hellmuth, 2009; Farghaly, 2010; Al Suwaiyan, 2018). Specifically, that speakers from the Maghreb region may face difficulties when communicating with speakers from the Mashriqi region (Horesh & Cotter, 2016; Medhat-Lecocq, 2017); and to solve this, language policy states that Modern Standard Arabic should always be used in these situations (Versteegh, 2014). However, acceptance of colloquial Arabic's contradicts the requirement for Modern Standard Arabic to be the solo variety of Arabic used, (this issue is explored in greater depth later in section 2.3.3 in this literature review and in

chapters 3, 4 & 5). Furthermore, the other issue with the Arab league's language

policy is the rejection of minority languages and communities from engaging with the

organisation itself (Pool, 1991; Spolsky, 2011; Or, 2017). Whilst the Arab league

presents itself as an organisation which represents multiple ethnicities and cultures

without discrimination between them, it is difficult to view the Ausbaucentric language

policy as inclusive as suggested.

2.3.2.4.4 African Union

There are two criticisms of the AU's language policy that were identified. Firstly, the

dependency on using colonial languages in a postcolonial era (Lin & Martin, 2005;

Shin & Kubota, 2010). Across the African Nations there have been social efforts to

decolonise the nation state in favour of developing an independent society which is

no longer impacted by the injustices the indigenous population faced under the

colonial rule (Wiliams, 2011; Juffermans, 2015). For the fact that most of the

languages within the AU's policy being representative of that historical colonial past,

can be considered an additional remnant of the colonial era that is yet to be adapted

or altered, however it is worth noting that this criticism is reduced since Swahili was

included as a working language. However, prior to the inclusion of Swahili, it could be

argued that the language policy encouraged colonial languages to remain prestigious

and defacto across the continent, as the use of the African languages may have

been supported by Article 11 (Kiyaga-Mulindwa, 1980; Jufferman, 2015), but the

social boon that is attributed to an official and working language was not attributed to

an African language. There are continual efforts to include further African-based

languages in the AU's language policy, through a movement to decolonise the

continent (Lin & Martin, 2005; Khepera, 2020). The second criticism is the rationale

presented for the maintained inclusion of colonial languages, which is not the

historical connections or the size of the speaking populations for the languages within

the continent of Africa, but rather the economic benefits that the population can enjoy

by being speakers of the colonial language (Ager, 2001). This presentation of

economic wealth and prosperity being connected to predominantly European

languages is further solidifying the colonial languages as having power and influence

in terms of social status in the African continent (Samuelson & Freedman, 2010). As

such the case that there may be economic benefits on a global scale to using the

languages selected in the AU's policy, particularly that they are, in part, global lingua

franca's (Williams, 2011; Kaplan, 2018), does so at the expense of encouraging

minority languages from the target population that the organisation represents, and

does not encourage global diversity and support of a minority language .

2.3.2.5 Building upon policies

Despite the criticisms identified for the current international language policies, it is

important to note that when developing language policy for disaster mitigation, that

resources available are limited (Quarantelli, 1991). This limitation is due to the nature

of disaster management itself, whereby the policies created are done so before an

event occurs; and there is always doubt towards the relevance of the policy, given

that it may in fact never be used (Carter, 2008). Therefore, to advocate for the

development of an entire language policy from scratch would be a difficult approach

in terms of practicalities, as funding bodies are unlikely to provide sufficient resources

for formulating and developing brand-new policies for a disaster communication

(Arai, 2013a; O'Brien et al., 2018). In part, the issue is the lack of certainty that the disaster will occur, with governments favouring the allocation of resources to current issues as opposed to future events (ibid).

From the four international language policies explored in this subchapter, there were shared positive factors which can support the adaption of the policies to suit the requirement for a disaster language policy. The shared features were: the identification of language as a human right; the importance of educating others in the languages' selected (to develop a wider speaking community); and the support of developing further language opportunities for the general populations, to increase the cultural and ethnic diversity within their respective nations. These features align with general approaches to disaster management and responses, those of maintaining the integrity of language rights as fundamental freedoms as well as the supporting the cultural identities of the speech community (Skutnabb-Kangas, 2006; Lindell, 2013; De Varennes, 2021). The UNLP, EULP and the AU all have overtly decreed the importance of approaching language differences with a positive attitude, with the value of linguistic and cultural diversity being held high by default. The AL encourages the localised Arabic's in the Arabic speaking nations, which highlights that intra-language variation is also not discouraged, like the other three organisations and policies outlined. Furthermore, the AU also supports the unification of the Arabic speaking communities as a single people (Hachimi, 2013), with a shared socio-ethnic identity by virtue of being an Arabic speaker and a nationality from an Arabic speaking nation (Horesh & Cotter, 2013; Versteegh, 2014).

There are positive shared characteristics which are relevant to the development of a disaster language policy within the framework of this thesis. Each of the organisations explored developed infrastructure for the enactment of their respective language policies. There are translation and interpretive services available for use in all of the respective organisations (Bassnett, 2013), which produce correspondence, research, and other documentation in all of the languages selected, such as the EU producing resolutions in all 24 languages (Ivana, 2020). Therefore, there are standardised language conventions for the 30+ languages identified across the four policies.

Furthermore, all four organisations examined have allocated resources to the development of teaching resources, for the languages included in their language policies, resulting in wide-ranging resources (Baldauf & Kaplan, 2005; Ferguson, 2006; Hult, 2012), including the training of interpreters (d'Almedia & Otcu-Grillman, 2013), and foreign language teachers (Shin & Kubota, 2010; Gorter & Cenoz, 2017), and the building of publication infrastructure for the printing and distribution of language material (Wright, 2016).

Whilst this research gap is significant, there are remaining issues that TWB (2017a) reported, which cannot be explained by top-down language policy alone, as such this literature review will continue. The following section will explore the status of socio-linguistic variation used within a language, to ascertain whether the communication barriers, of Arabic speakers being unable to understand one-another, experienced by migrants, as reported in TWB (2017a), was in part a result of language variation within Arabic. The variation avenue is motivated by the diversity of language, with

both the UN and the EU governing bodies highlighting how differences between

language speakers can become an issue when providing translative support in formal

discussions. Furthermore, language use in the Arabic League, which is a

monolingual organisation using Arabic as the solo language for discourse, has been

described as limited in context when there are speakers from across the Arabic

speaking regions (Or, 2017). Therefore, it is not presumptuous to consider language

variation, particularly of Arabic, as a factor that may explain, in part, the status quo of

communication difficulties within existing international language policies.

2.3.3 The sociolinguistics of Arabic: an overview

Likewise, colloquial Arabics have also experienced language change, however the

changes have been diverse and widespread so this thesis will not detail the specific

changes noted across the Arabic speaking nations. There are two primary categories

of colloquial Arabic's, there is the Maghrebi which is the varieties spoken across the

North-Western African nations (Hachimi, 2013); and there is the Mashriqi which is the

varieties spoken across the north-eastern and Middle East nations (Zbib, et al., 2012;

Zaiden and Calliuson-Burch, 2014) Moreover, the TWB (2017a) report focused on

refugees fleeing Syria, which is a nation that falls into the Mashriqi Arabic region, as

such colloquial Arabic, and the changes that occurred to them, will be focused in this

thesis on Mashriqi varieties. This is motivated by Mashriqi Arabic's being grouped

together by shared history and evidenced semblance of linguistic similarity (Akbacak

et al., 2011; Albrini et al., 2011; Abushariah et al., 2012; Kwaik et al., (2018b).

The agreement as to this two-way distinction is not universally accepted across the

literature, in fact there are efforts to expand the number of groups further than a

simple binary choice. Zhib et al. (2012) proposed that there are at least five groups across the Northern African and middle eastern nations, including the separation of Mashriqi into Egyptian (EA), Gulf (GA) and Levantine (LA). Abdul-Mageed et al., (2018) further proposed that Arabic be split into smaller groups, such as separating Gulf Arabic into multiple distinctive varieties. Additionally, the language database Ethnologue (2022) also supports the expansion of Arabic classification by recording multiple Arabic varieties as individual languages, such as Egyptian QA, and Levantine QA. The motivating data for the reclassification of Arabic is linguistic distance, comprehension, shared identity and/or intelligibility; for instance, Zhib (2014) used machine translation, and the level of issues in translating between different Arabic's as a primary factor in classifying Arabic varieties.

2.3.3.1 Linguistic Variation

Linguistic variation within languages is a well acknowledged phenomena for languages (Horesh & Cotter, 2016), a regular process of change for the grammatical and functional characteristics of the language's system (Walters,1996; Almujaiwel, 2020). The phenomenon of language contact is also attributed to the rate and direction of language variation and change (Albirini, 2016), as when there are two (or more) languages or linguistic varieties bordering one another, both will have an influence on the other (Manfredi & Tosco, 2018). The influence can be any language feature, or any part therein, such as the phonological units (Mustafawi, 2017) or the lexicon (Almujaiwel, 2020).

Language contact occurs when there are two or more distinct linguistic varieties in engagement across a single environment (Trudgill, 2009), this can be between two

distinct languages, or it can also be between two varieties that are considered part of

the same language (Embarki, et al., 2007). Regardless of whether the language

variety in question has received the prestige status to be recognised as a separate

language (Tosco, 2008), the processes that change the linguistic systems are similar,

such as phonological levelling (Al-Royaie, 2013) or lexical truncation (Parkinson,

1991).

Language variation is used as a tool for demarking the differences between unique

systems of linguistic communication, such as distinguishing between multiple

languages (Al-Jallad, 2019; Simon & Lewis, 2020; Tamburelli, 2021). The

demarcation process produces isographs, which signify where the border between

the two unique varieties exists (Theodoropoulou & Tyler, 2014, Towler, 2018). These

borders are commonly geographical, as it allows identification of the regional space

that the unique variety is prevalent or present in. This approach is widely used to

explain the differences between dialects and locations across multiple nations, for

instance in Italy it is used to demark the cultural boundaries (Tamburelli, 2014),

whereas in India it is used in part with the caste system in the northern territories

(Mohanty, 2019). Language identity is connected to language variation (Bassouiney,

2014), which is in turn connected to the social framework of the community (ibid); the

complicated relationship between language change and societal change is

sociolinguistics and political. When there is a political shift in belief or ideas, the

language used across the area typically changes to reflect the beliefs that are

changing (Theodoropoulou, 2015; Omar & Ilyas, 2018); this includes the positive

beliefs of those who want to change, and thus encourage change, as well as those

resistant to change, who are more conservative. As a society develops, the language adapts, and with it the language reflects the current and the historical changes in social norms (Owens, 2015). This process of a language changing to reflect the current situation and the demands of the current population is universal (de Jong, 2004). As such, language variation itself is equally as universal, as language is ever changing to fit the needs of the population itself (Cotter, 2017).

An important distinction to clarify here is the difference between a natural language and taught. A natural language is when a linguistic system has developed through predominantly use, and as such, experiences a high level of fluctuation changes and innovation (Ibrahim, 2021). In contrast taught languages are ones which are not naturally spoken (Alresaini, 2016), in that they are taught rather than acquired, examples of taught languages include Modern Standard Arabic which is a standardised version of a written language that is then being applied to a population through school (Versteegh, 2014). This distinction is important to remember as this thesis will be looking into the relationship between natural linguistic systems and taught linguistic systems, particularly in the realm of Arabic; a language which has a distinction between a natural system (with colloquial Arabics (QA)) and a taught system (with standard Arabics (SA)) in use at the same time in the population. Before discussing the complicated situation that is Arabic sociolinguistics and the relationship between the natural linguist existence and the taught, it is first important to briefly overview the changes that have resulted in the current status quo. Specifically, how Arabic has developed into holding both natural and taught linguistic systems simultaneously.

Classical Arabic (CA) is the language which the Quran is written in, as such it is an important that Muslims understand the language itself and can comprehend the nuances described in the text (Chejne, 1968). The Quran has been a major source of linguistic policy for the Arabic speaking communities for over a thousand years, and so the classical Arabic as it is known, from the Quran has been used as a bedrock for marking correct and incorrect Arabic (Cadora, 1989). Alongside classical Arabic, which was historical used as the defacto written standard (Owens, 2015), there was also the local colloquial Arabic for the day-to-day experiences (Farghaly, 2010). This included family communications, commerce discourses and in between all other spoken situations. As such there were (at least) two distinctive and unique language varieties in the Arabic speaking communities simultaneously, this was a status quo pre-modernisation of Arabic and so the consideration of pre-classical Arabic is a step too far back in history to retain relevance to the modern socialistic framework of today (Versteegh, 2014).

The development of the printing press, the Internet and the information age resulted in classical Arabic no longer being considered fit for purpose across the Arabic speaking regions (Al-Sobh et al., 2015). It is not to say that classical Arabic was diminished in terms of being correct or perceived as being the pure form of Arabic, rather than the actual lexicon itself was insufficient to describe the new technologies and situations that were being experienced by the population. When this occurred, during the earlier days of modernisation, the population were forced to use either the colloquial Arabic to describe what was occurring or to use a loan word from a foreign language such as English or French which were the lingua francas' during the

industrial era (Kwaik et al., 2018b). To counteract this, the Arabic language

academies wrote in unison on modernising the written standard Arabic of the time,

which is classical Arabic, for the new era (Qafisheh, 1974; Horesh & Cotter, 2017).

This process included identifying methods and approaches that should be used when

considering the inclusion of word or a borrowed word, such as the extent that

loanwords would experience arabification; specifically, on how the loaned items and

words would be adapted and realised in an Arabic medium discourse (Saiegh-

Haddad, 2004). When adapting classical Arabic and providing guidance on the

introduction of loanwords, it is important to recognise that this process was not strictly

formal in terms of legal process or absolutism; and there was no agreed-upon

guidebook or instruction manual on the changes being made to Arabic in the

development of modern standard (Albirini, 2011). Moreover, it was the combination of

efforts by a multitude of language academies and political organisations which

through a combined general approach formed modern standard (Khalil, 2012).

A lack of a coordinated single point approach to unifying Arabic for the modernisation

of the language, resulted in variation between the modern standard in use today.

Whilst it is generally considered that modern standard follows a consistent approach

in terms of functionality and grammatical units, the same cannot be said about the

spoken realisation of modern standard (Albirini, 2014; Bentley, 2015). The written

elements of modern standard are heavily regularised (Versteegh, 2014), given that

the route for the realisation of written standard Arabic lies directly from classical

Arabic in the literature of the Quran, (Holes, 2005) whereas, spoken realisation of

modern standard was impacted and influenced by the colloquial Arabic's in use

around each language academy or political sphere that were involved in the

modernisation process (Elmahdy et al., 2009; Almujaiwel 2020).

In the case of Arabic, there has been historical resistance to variation by puritans,

who have often argued that there is a pure Arabic and that variation from their

proposed original Arabic is dirty and wrong (Cadora, 1989); like the situation in

France where other variants and languages are formally classified as dirty French

(Lodge, 1991; Judge, 2006). There is much debate into which Arabic is to be

considered the original and pure Arabic, for instance Walters (1996) argues that a

variant from pre-Islamic Arabic should be considered the pure variant, since they are

the precursor to modern Arabic; an argument that is rooted in the evidence of

variation from pre-Islamic dialects, such as the use of the epenthetic vowel in

Western dialects whereas the East did not have final consonant clusters with a

vowel, e.g. kalima/kilma 'word', unuq/unq 'neck' (Bassiouney, 2014; Horesh & Cotter,

2016). However, there are grammarians who argue that post-Mohammed is the era

in which the pure Arabic should be considered from, as there is a link between the

purity belief and the religious beliefs, as the Quran states that the Arabic used within

it is the correct from of the language (Trudgil, 1992). As a result, the Bedouin tribes of

the desert regions were considered, and in some extents are still being considered,

the defacto authority on the Arabic language (Versteegh, 2014). This is due to two

different reasons, firstly the prophet Mohammed was from the desert-nomadic tribe

culture, so his idiolectic linguistic features were similar to the Bedouin tribes, thus

there was a link made between the religion and the tribal linguistic features

(Versteegh, 1993). Therefore, the desert linguistic features became the preferred in

usage within the cities and metropolitan regions (Miller, 2007). The second reason justifying the Bedouin, as the defacto leaders of linguistic policy until the late 20th century, is due to the conservative nature of the community (Cadora, 1989), because they are isolated from the metropolitan areas there has been less linguistic innovation (Parkinson, 1991; Mir, 2019). The result of the conservatism is that the language use of modern Bedouin is similar to the language use within the time in which Mohammed roamed the earth. Historically, the Bedouin tribes of the desert regions were considered the keepers of perfect Arabic, the retainers of original faith since the arrival of Islam in the Arab-world, which provided the Bedouin tribes with immense social prestige and influence (Versteegh, 1993). There are recorded examples of the Bedouin tribes testing scholars of the Quran and Arabic by examining their oral language use, and correcting the errors made, errors which in other cases would be considered naturally occurring variants (Ownens, 2015). Hyper-correction is a commonly recorded feature of the Bedouin tribes historically and in some areas of the Arab world the phenomena is ubiquitous with the Bedouin tribes (Parkinson, 1991).

An additional factor to account for in the modernisation of Arabic, is that of the efforts of resistance to change. Particularly, those which were motived by religious ideologies or by religious figures. The resistances were linked to the religious prestige of classical Arabic, which is considered as connected to the original language of the Quran (Bassiouney, 2009). The debate of modernising Arabic focused on the allowance of international languages as origins for words, for borrowings and source material, with the conservative actors and organisations

favouring an avoidance of using foreign language words in all but essential changes (Albirini et al., 2011). Alternatively, there was also an effort for the adoption of loanwords, but with modifications; to ensure that the words became standardised in the phonotactic and morphological rules of classical Arabic. This was a compromise between accepting neologisms and borrowings into standard Arabic, and the retainment of the classical Arabic routes; in an effort to retain social cohesion between the religious and the secular factions across the Arabic speaking nations (Versteegh, 2014). The connection of classical Arabic and the religious prestige of Islam have been reinforced by de facto language policy for over a millennium.

Since the widespread deployment of MSA, classical Arabic has undergone a process of functionality change. Modern Standard Arabic is now the de facto Arabic variety that is used in political discourses, educational contexts and in the running of the state (Albirini, 2016). Whereas classical Arabic has reduced in functionality to being predominantly restricted to religious context and environments, such as in mosques and during public holidays (Bassouiney, 2014). This change in functionality between the two standard Arabics is a well-documented feature of diglossia (Ferguson, 1959; Fishman, 1968), which is when there are two language varieties in a single community and each variety has a specific context of usage, and as such there is a sociolinguistic disparity both the language varieties (An aspect which will be investigated in depth later in *2.3.3.3 Arabic Diglossia*). However, the key point to consider is how standard Arabic, which has been described as immovable (Djennane, 2014), has experienced rapid change in the 50 years (Ibrahim, 2021).

Within the general Mashriqi group, there are noticeable linguistic differences, both in terms of the language system itself and the realisation of Arabic overall (Amayreh & Dyson, 1998; Amayrah, 2003; Miller, 2005; Alresaini, 2016; Kwaik et al., 2018a, 2018b). There are syntactic differences between the SAs and the QAs, predominantly the sentence pattern; as SA uses a verb-subject-object, whereas the QA's use subject-verb-object (Hamam, 2014; Versteegh, 2014). Additionally, there are phonological differences in inventories, for instance, the phoneme [q] can be realised differently between Gulf Arabic and Egyptian Arabic; with /q/, /dʒ/ and /g/ in Gulf (Holes, 1990), whereas Egyptian also has /ʔ/ included (Watson, 2007). Another phonological example is [j] which is realised as /j/ or /dʒ/ in Gulf, however in Egyptian it can be realised as both sounds as Gulf plus /g/, /d/, /gɤ/ or /jˀ/ (Amayrah, 2003). Furthermore, it is important to recognise the impact that phonological differences can have on the realisation of MSA, given that MSA production is reliant on the QA of the speaker (Albrini, 2014; Asadi & Abu-Rabia, 2021). Egyptian speakers, when using MSA, will realise speech sounds differently to Gulf speakers, and these differences can result in a reduction of intelligibility and comprehension, as the Egyptian patterns are anomalous in Gulf (Horesh & Cotter, 2016). In addition, there are lexical differences, that are rooted in the historical borrowings that each QA community adopted historically; which when combined with the morphological differences produce variation across the Arabic continuum (Bassiouney, 2009; Watson, 2002, Versteegh, 2014; Albrini, 2016).

For instance, if the speaker were from the northern region of Syria and they spoke to members of the Egyptian community, the Syrian would be directed to join the Syrian

community as it is more suitable to where they belong. There are of course exceptions to this trend, as refugees can find other unifiers besides linguistic characteristics, such as orphans who have no carer often band together as their family state similarity supersedes the linguistic differences (Sourander, 2003). Overall, linguistic differences are a primary divider and unifier of refugees in the aid camps.

The overarching issue in the Southern European Refugee Crisis is the lack of available interpreters, as only 1/13 got access to an interpreter at some point in the processing of the refugees into the host nation (TWB, 2017a). It is important to note that this figure is only referring to when the refugees get an interpreter at any point in the process; the number of refugees getting access to an interpreter throughout the entire process, from joining the refugee community to relocation in a new host country, is far smaller and has not been accounted for in studies of refugees (Gerver, 2021b). This omission is a result of a research approach that aims to increase the accessibility first, so that in the future more consistent interpretations services can be provided, i.e., the start point is ensuring that refugees get sufficient language support initially, and once this is established, then focusing on ensuring the support remains stable (O'Brien & Federici, 2020). There is an additional issue within the matching of refugees to appropriate language support; that of whether the refugees can understand the language. For example, TWB (2017a) reported that Arabic language support was inaccessible to some Arabic speakers, due to the linguistic differences between Arabic varieties, yet the interpreters were allocated on the basis that they speak the same language as the refugees.

2.3.3.2 Arabic Ausbaucentrism

Arabic has been described as linguistically diverse in terms of dialectical variation, yet this description is in-fact easier described by Arabic being Ausbaucentric. This classification of Arabic language policies and demarcation as suffering from Ausbau-focused approaches is motivated from the widespread acceptance of intelligibility differences between Arabic varieties, whilst also stating that each of the Arabic varieties are all part of the same language. This rejection of demarking language varieties as different from the 'umbrella' language is a key feature of Ausbaucentrism (as discussed in section 2.3.1.4 Abstand and Ausbau). Alongside Arabic being Ausbaucentric, there is an establishing intra-Ausbau language state of diglossia, although this is often ignored within the context of the Ausbau-Abstand, predominantly, due to all Arabic varieties being classified as part of a single language, rather than a collection of languages within a single overarching umbrella language.

2.3.3.3 Arabic Diglossia

'Most [languages] are far less diglossic than Arabic' (Kaye, 2001: pp. 118). Arabic is diverse enough that Arabic speakers from different backgrounds can struggle understanding media broadcasts in another Arabic variety (Kaye, 2001); this supports the idea that using mutual intelligibility is not a 'good gauge' (pp. 124) for demarking a language community. It is important to identify with Kaye's (ibid) point, which is that when explaining the differences between two dialects of the same language using mutual intelligibility, the suitability of the result is questionable.

However, a criticism of this view is the Ausbaucentric bias underlying the argument. If the method of mutual intelligibility does not support the consideration of two or more varieties as being part of the same language, then the issue underlying the combining of both varieties into one entity is that communication, as in the ability to actual convey meaning in discourses, is lost. This paradoxical argument against mutual intelligibility is often rooted in the ideology that the two linguistic varieties under investigation are already known as being part of the same language, as in the case of Kaye (2001), which is not necessarily the case from a communication perspective.

Ferguson (1959) proposed that diglossic varieties are realised in a uniform manner, specifically that the production of the standardised H variety would be consistence across the communities, whereas the L varieties are inconsistent, as a result of their colloquial status, with variation in realisation being commonplace. Kaye (2001) highlights how MSA does not fit with Ferguson's (1959) model, particularly in relation to phonological realisation and pronunciation of MSA. There is no standard MSA accent or speaking system that is universal across the Arabic speaking nations, in fact, the MSA is influenced by the QA of the speakers and the speech environments. (Albirini, 2014) As a result, the current situation of Arabic indicates that MSA speech is Egypt will be closer to Egyptian QA than MSA in Saudi Arabic (when Gulf Arabic is the QA) (Bassouniney 2009). Arabic diglossia is a continuum, with the influence of the QA on MSA being a primary factor in the variations' realisation and application (Saiegh-Haddad, 2007).

Historical influences are a factor in the development of the modern QA's, as well as with the realisation of MSA across the Arabic colloquial groups, particularly the Maghrebi and the Mashriqi (Versteegh, 2014). The Maghreb colloquial Arabics were influenced by the French colonial empire, with QA's borrowing lexical items from the French language itself whilst also undergoing phonological and morphological change due to the language contact and influence (Horesh & Cotter, 2017). QA's, such as Algerian and Moroccan, in modern day, have retained features from the colonial era in the language use (Bassouiney, 2009). The French-influenced QA's (Maghreb) contrast with the British-influenced areas (Mashriq) (Hachimi, 2013). In the areas that were under British colonial rule, the predominant influence for language change of QA's was taken from English, in a similar fashion to the French dominions (Versteegh, 2014). English words were loaned, and there were changes to the QA's of the dominions through the influence and language contact (Walters, 1996). The historical roots of political control and colonialisation are the primary reason as to the east-west distinction in Arabic (Versteegh, 2014). In modern day, the divide is significant, as there are Arabic communities that are deemed too far away, both linguistically (Bassouiney, 2009; Versteegh, 2014; Horesh & Cotter, 2018) and socially (Hachimi, 2013), from their latitudinal neighbours. A popularised example is the difference between Dirija Arabic (Morrocan QA) and Najdi Arabic (Saudi Arabian QA), as it is widely documented that there is little intelligibility between the two Arabic's (Bassouiney, 2014), to the extent that it has been recorded as a language barrier (Zbib et al., 2012; Bidaoui, 2016).

Currently, Egyptian QA, the L variety, has expanded in use, and has gained prestige in political, social, and legal contexts (Sohb et al., 2015; Ibrahim, 2021); this has been attributed to the increase in general education, and the use of EA as the medium for teaching in main-stream Egyptian schools (Snow, 2013). Traditional diglossia with Arabic is characterised as being integrally linked to the statehood and religion of the state, Versteegh (2014) highlighted that the nationalism of Arabic culture was connected to the efforts to preserve the status of CA, as a pseudo-H variety; furthermore, the level of education across the state, historically, provided a stable backdrop to prevent the development of EA beyond the L-status, by virtue of the power; politically, social and militarily, being retained solely by an established elite.

Ready (2018) further agrees with Snow's (2013) proposition that national identity is connected to diglossia, stating that the 'concept of diglossia… is a key component for national development' (pp. 177). Additionally, the concept that language identity, and social status is connected to national development is not nuanced (Spolsky, 2011; Wright, 2016); however, the application of national development in established diglossic states, where it is widely accepted that communication issues exist (due to mutual unintelligibility), is more nuanced. This is especially relevant given that Arabic states are under-going linguistic and social change at faster rates than ever before, predominantly due to technological advancements and the development of social media; for instance, Egyptian Standard Arabic is becoming more acceptable in written form, which is considered a result of the modernisation and technological development of media and telecommunications (Khalil, 2012; Alruily, 2020).

Furthermore, Snow (2013) stated that Arabic bucked the diglossic trend, as the language had not experienced typical change, i.e., the reduction in prestige or usage of the H variety in favour of the L variety (Ferguson, 1991; Djennane, 2014). Yet, Saiegh-Haddad & Lina's (2018) position was that change was continuous, and the use of SA, whilst still being prolific, was lessening in favour of QA's which were historically limited to L status. It is important to recognise that it is individuals in the Arabic speaking nations that need to be proficient in both MSA and the local QA to be able to function uninhibited in the local community. So, in the limited and simplified sense of the linguistic background, there has been little change in diglossia, as both a SA and a QA are needed, and that a SA is used for formal context and QA are for less-formal context; however, the status quo is less stable than Snow (2013) observed.

Within the Arabic language diglossic continuum is a set of higher L varieties amongst a larger set of lower L varieties (Kaye, 2001). The distinction of the H and L varieties varying by nation or major geographical boundaries (Al Suwaiyan, 2018). For instance, the highest L variety of Egypt is Cairene Arabic, which is the QA spoken and attributed to the capital city Cairo (Watson, 2007). Cairene does not appear within the neighbouring nation of Saudi Arabia as a higher L variety, rather so, as a lower L variety, as Gulf Arabic is the highest L variety in Saudi Arabia (Omar & Ilyas, 2018). The MSA produced by any nation is influenced by the features of the highest L variety, which typically is the QA spoken in the regional capital; as the socio-economic advantages for using and learning are greatest for the individual in the nation (Ager, 2001; Trudgill, 2009; Wright, 2016), additionally, the MSA is regulated

by the government, which typically use the highest L variety in the capital for the business of the state and government. For instance, the MSA of Egypt is influenced by Cairene Arabic, since the Cairene is the highest QA in Egypt, and is the de facto language variety in Cairo (Bassiouney, 2014).

When comparing language use and language variety status's cross-nationally, it is important to provide a method of comparison that matches the varieties investigated in equal conditions. For instance, if an L variety of Arabic is compared to the H variety of MSA, and the research is assessing the general attitudes towards the two varieties, then the result could be biased towards the more prestigious H variety. This issue has been highlighted across language attitudes and beliefs studies previously (Holes, 2011; Abdel-Rahman, 2016; Ready, 2018); with it understood and agreed across socio-linguistics that a society will favour supporting the H variety in attitude testing, especially when the measurements are asking questions which are connected to a value of social attitude explicitly; for instance, when a question asks whether a QA variety should be considered as an equal to the MSA variety presented; the responses will reflect the overall socially expected attitude, not necessarily the attitude of the individual.

To compare the language varieties whilst limiting any potential diglossic related bias, a balanced set of language varieties should be used, with all varieties including being of equal status within their home-nation. This will allow for clearer comparisons between the Arabic varieties in terms of intelligibility, as well as allowing for implicit attitudes to be solicited without overarching issues of underlying bias. In this thesis, it is proposed that the QA's that should be used in investigating intelligibility should be

those found as de facto variety in a national capital of an Arabic speaking nation; as this will ensure that all QA's involved will benefit from equal status and prestige overall, as each would be prestigious enough to have influence the localised realisation of MSA (Watson, 2007; Albirini, 2014). Furthermore, it is important to account for additional prestige that can be attributed from the socio-economic power or status of the nation from which each QA is used in. For instance, a comparison between Cairo and Bamako would have to account for the imbalance in power between the two QA's populations, as the Egyptians are more affluent due to their nation benefiting from stable trade networks via the Suez Canal (Bassiouney, 2009), whereas Mali, is land-locked and benefits less from cross-border trade (ibid). Contrastively, a comparison between Cairo and Riyadh is more suitable, as both cities are hubs of finances, trade, and culture; and their respective nations are affluent and influential both regionally and globally (Watson, 2007; Versteegh, 2014); also, the QA's of each have comparable speaker populations. It is impossible to have truly equal comparison available, but equity in comparisons is needed to retain rigor of results and conclusions.

2.3.3.4 Overview and next steps

This subchapter has explored the concepts of diglossia and bilingualism, both as theory and as applied to a language, in this case Arabic. From this exploration, the status of Arabic has been highlighted as unique, as it has maintained stable diglossia and bilingualism in an equilibrium without intentional language planning being presented or conducted to retain this status. This sub-chapter has also investigated

the relationships between the diglossia of colloquial Arabics, both with the standard varieties of Arabic and other comparable colloquial Arabics; with a research gap of Arabic variation and demarcation being identified due to a lack of research into the status of modern colloquial Arabics on an international diglossia community. The prevalence of Egyptian Arabic has impacted the language use across mainstream media in the Arabic speaking regions, however, the extent to which this has changed the diglossic status of other nations is unknown; resulting in a missing topic of how Egyptian Arabic has permeated into other Arabic cultures, enough to become considered equal in the community to the localised colloquial Arabic. A further significant issue in Arabic that has been identified is that there can be communication issues between Arabic speakers themselves, as a result of linguistics differences, which echo's the TWB (2017a) migrants report. The extent to which Arabic speakers can understand each other in general discourse is disputed, with the rational for this situation potentially laying fault with the demarcation of the Arabic language itself, and an overapplication of diglossia and Ausbau (Manfredi & Tosco, 2018). If there are issues in peace, then when the context is stressed, and representative of the situations reported in TWB (2017a), the negative impacts on communication are feasibly going to be experienced more severely or realised with greater furore. Current research has not provided answers, explanations, or investigations into the status of Arabic language use in emergency contexts. At this stage, this thesis proposes that this gap should be a focus point for further research, and to explore this there will be an investigation on the status quo of measuring communication,

intelligibility and comprehension is warranted, to assess whether the answers can be

identified from existing research.

2.3.4 Intelligibility

2.3.4.1 Defining Intelligibility

A simplified definition of intelligibility is whether a message can be understood by

another human being (Flipsen, 2006). This encapsulates the essence of the general

working definition of intelligibility in linguistics, since understanding of message is a

core aspect of intelligibility. This definition, which is effective in being a simplification,

is crude; and gives rise to two issues. Firstly, when measuring intelligibility, and what

can be classified as being intelligible, how is, or should, understanding be recorded.

The simplification implies that the answers to intelligibility between two speakers can

be answered using a binary *yes or no* design, yet communication accuracy and

understanding are not absolute. As Gooskens (2017) argued, the intelligibility of a

message is a gradient factor, in that the understandability should be considered as a

continuum.

The second issue relates to which perspective understandability is measured from,

the speaker or the listener. If it is the former, the intelligibility, or perceived

understandability measurements can inform whether a speaker feels that they are

understandable (Fuse et al., 2018). Whereas, if it is the listener, the measurements

can inform whether a message can itself be understood and processed correctly as

intended by the speaker or will an altered or limited message be retrieved

(Kuperberg & Jaeger, 2016). Additionally, consider the situations where the

speaker's perceived intelligibility is different to the message understood by the listener, both measurements would contradict one-another, leading to a disputed sense of intelligibility. This contrast brings the final rules of what is intelligible into dispute. Simplified definitions are useful in outlining the basic and general concept of intelligibility as *to be understood* (Flipsen, 2006), but there are multiple additional facets that need to be considered with intelligibility. At this stage, there are two facets of intelligibility to consider, firstly the ability to understand, and secondly the inverse, that of when there is no intelligibility.

It is important to recognise that there is a distinction between written intelligibility of language and spoken language, specifically relating to the different features that could influence or impact intelligibility, that are unique to each medium of communication (McCloy et al., 2015). A primary distinctive feature is accent, which is found in spoken language exclusively, as a by-product of the production of the acoustics of spoken language (Munro & Derwing, 2011; Ngo et al., 2020). Another distinction is that the error rates between both mediums are different, with spoken featuring greater error rates than written (Perlmutter, 1989); as it is not uncommon for a speaker to mis-produce a target sound, by incorrectly realising a phoneme, therefore producing an unintended speech sound (Munro, 1993). Written language is more structured and regulated, due to standardisation, whereas spoken is descriptive and spontaneous in most discourses (Smith & Nelson, 2019). Therefore, when investigating intelligibility, the targeted medium of communication needs to be identified, with research considered from studies which use said medium. This avoids overlap of results and discussions that are non-comparable, by virtue of the features

involved. From the motivation, there were issues for both spoken and written communication, however, the focus will be placed on the actions of the interpretive services, as opposed to the translation services, thus the thesis prioritises focus and considerations on spoken rather than written language when investigating intelligibility.

A significant issue in intelligibility research is found in the defining of the phenomenon itself. There is inconsistency and a lack of agreement as to what the term itself consists of and includes. Across multiple domains, from speech pathology to classroom pedagogy, there are differing working definitions, as such, the following section will examine the different working uses of the term intelligibility, with the aim of producing a working definition for this thesis moving forward.

Nelson (1982) proposed that for intelligibility to be considered, a vital element was whether the correct intention was understood by the listener, in that the nuances from the speaker were accurately received and decoded by the listener, a view supported by Gass & Varonis (1984), who investigated the accuracy of transcribing spoken word with second language learners, with the focus on intelligibility being directly related to the replication of the intended meaning of the listened sentence. An approach replicated by Munro & Derwing (1995), who instructed native speakers to transcribe utterances of second language learners speech, to investigate their definition, which can be summarised as, the extent to which a native speaker can understand an intended meaning in a message. Whilst the above studies were produced with the context of learning a second language, with the focus being on the way in which a speaker can be considered proficient in listening comprehension,

there are similarities with another element of linguistics, that of translations. Similar to equivalency in translation studies, where the pragmatic, and non-literal, meaning of a statement or message needs to be reflected in total in another language (Bassnett, 2013). From this, the extent to which a message can be understood, with any intended nuances, will be included for the consideration of intelligibility in relation to the understanding of the listener in spoken discourse.

Within the realm of speech pathology, research favours defining intelligibility through spoken medium more so that written. For instance, Yorkston et al, (1996), proposed that intelligibility is linked with the spoken aspect of language, with a definition that focused on the extent that a listener can understand an acoustic signal. This definition has been adopted by future research conducted, such as Miller (2013) and McCloy et al (2015), which both investigated the production of speech and the understanding of language with children with speech sound disorders.

However, in contrast, Gooskens et al (2010) expanded the definition of intelligibility in spoken language, to 'the ensemble of properties that allow a native listener of a language to correctly recognize the linguistic units… produced by the speaker' (pp.1022). Furthermore, Gooskens et al (2010) distinguishes intelligibility, as an ability to recognize linguistic units, from comprehension, which is the ability to comprehend the meaning of a statement. This approach supports the notion of speech understanding, as proposed by van Heuven (2008), where the measuring of comprehension is based on the extent to which the meaning of a statement is understood with accuracy. This splitting of factors, allows for the identification of a language to be separate from the ability to understand the messaging in said

language, allowing for the situations when a person can understand the identity of a speaker, but they are not competent in their ability to decode beyond this step (Nelson, 1982). A conceptualisation for this can be found in the earlier stages of learning a second language, where part of the process for developing spoken and listening proficiency is the ability to identify if the language that is being learned is being heard (Munro & Derwing, 1999); whereas the later stages of language learning focus on the ability to understand the messaging, to facilitate discussion and provide a relevant reply in the target second language (ibid) .

Throughout the distinctions explored thus far, none have addressed the question of what the status of intelligibility, or comprehension, is when the differences are minor. The distinction between the ability to understand and to identify is based on the notion that neither are present from the onset of the investigation. The issue here is found when intelligibility is mutual, in the positive sense, whereby the communication between two individuals is understood, and identified in tandem without limitation or issues (Gooskens & van Heuven, 2021). Within mutual intelligibility, a distinction that differentiates the phenomenon from general intelligibility and comprehension is the focus on linguistic distance, as a measure of unification of a set of language varieties into a singular language (Tang & van Heuven, 2009). Early definitions for mutual intelligibility focused on the vice versa of intelligibility, such as Cheng's (1996) where the average intelligibility between two speakers was recorded in both directions of communication i.e., when the roles of speakers and listener where swapped. Goosken et al (2010) further summaries mutual intelligibility as to 'the degree to which the interactants are intelligible to one another' (pp.1022).

After this brief examination of the unanswerable questions that arise from a simplified definition, this thesis takes the approach that a catch-all definition for all of intelligibility is too limited and restrictive for academic use herein. As such, the following section will provide a set of working definitions for each sub-section, or divisional point, of intelligibility. This will booster the simplified definition of to be understood into a combination of targeted and precise working definitions for further use. There are four working definitions proposed, all of which are produced for research use in spoken language research:

– Unintelligibility – A situation when a message cannot be understood, recognised, or decoded between two or more individuals in a single spoken discourse event.

– Perceived intelligibility – the belief of ones' own, or another's, understanding of the linguistic features or intended meaning of either a heard statement from another speaker or in one's production of a spoken statement

– Functional intelligibility (herein consider this the default definition in this thesis for intelligibility) – the extent to which one can either understand or produce messaging, with the intended meaning, linguistic units and nuances retained for either oneself, as a listener, or another as the listener.

– Mutual intelligibility – the extent to which a language variety can be understood without inhibition or difficulty, both in terms of recognising the linguistic units and understanding the intended meaning of messaging, when the listener is using a language variety that has been classified as being different to the speaker's.

There are two primary groups that intelligibility research can be categorised into: the perceived and the functional. Perceived focuses on the beliefs of individuals in relation to how much they believe they understand or identify from spoken language use (Schüppert et al., 2015). Whereas the functional focuses on the precision of understanding, with comprehension being primarily from this approach (Gooskens & van Heuven, 2020). Both categories measure a part of the overall concept of intelligibility for a speech community, with the functional measuring the direct level of intelligibility or comprehension, whereas the perceptual recording the underlying attitudes and confidences. This two-way distinction is beneficial, as it separates the objective measurements from the subjective; whilst allowing for both qualitive data approaches to be used in tandem with quantitative data approaches, to produce a clearer and balanced insight into the status quo of intelligibility and comprehension in the investigated populations. When investigating mutual intelligibility, both aspects can be tested and measured, as the state of intelligibility, being as non-inhibitory due to a lack of distance between speakers, requires both the functional elements and the perceptual element being in alignment.

2.3.4.2 Why is intelligibility relevant to emergency language policies?

In a disaster, the local population will be at a high risk of harm; and the aim of disaster management is to alleviate this risk by conveying the situation to the population and supporting the population to reduce their individual risk of harm (Carter, 2008; Cadwell, 2021), such as the relocation from a hot-zone, or directions to shelters. This process requires the communication between the disaster management agencies and the impacted population to be clear, understandable, and

usable. If a population cannot understand any disaster warning messages, then they will not understand the situations' risk. As such, it is established across disaster management that accurate understanding between agencies and the impacted populations is not simply desirable, it is essential. For context, disaster management functions to reduce risk of harm, and the rejection of advice or instructions will at best maintain the level of risk; a level which indicates a likelihood of human injury. It is not an exaggeration to note that the rejection of disaster management orders can lead to aggravated levels of risk, in which the likelihood of survival decreases rapidly; this notion highlights that the context demands that communication is essential. Intelligibility, therefore, is an essential criterion for language policies for emergency situations and disaster management.

There is little research in the field of disaster communication and the intelligibility thereof, in fact, the majority of evidence found is based on Japan's tsunami warning systems, with the efforts focused on the impact of additional factors in the processing of public announcements. Hodoshima (2019), investigated the impact of speech urgency and text congruency in verbal disaster warnings, to assess whether the expectedness of the message, and the accurate perception of risk were connected to the intelligibility of the message. Finding that an increase in urgency, both in terms of production and in terms of perception, were related to higher intelligibility rates. However, Hodoshima (ibid) also identified that a multi-modal warning system must be conducted in tandem, with both messages being equivalent, as the intelligibility of messaging is severely reduced if the speech heard, and text read are not in total alignment. Other factors that impact intelligibility, in emergency contexts include

speaker gender and speaking rate; as (Ofuji & Ogasawara, 2018) investigated the effects of gender, fundamental frequencies and speaking rate on intelligibility of public announcements, particularly with female speakers. Finding that as speech rate increased, the perceived urgency also increased; and these two factors in combination improve intelligibility of disaster warning messages. Overall, both studies highlighted the need for additional research into the intelligibility of contextual based messaging for disaster warning systems and emergencies, as the situational factors in emergencies are vastly different. For instance, in an emergency, the environment is unstable, in contrast to peacetime environments; furthermore, the messaging's used in emergencies are often exclusive to the high-risk contexts and the lack of day-to-day use of these terms registers the severity in the listener (Sandman, 1989; 1993). To illustrate consider a warning message used by Ofuji & Ogasawara (2018), the phrase 'A big tsunami is coming. Please evacuate immediately' (pp. 57), is unexpected in general, peacetime situations (when there is low risk), and upon hearing the message, individuals register a sense of severity (or increase of risk) and thus focus on the intimate details in the message.

2.3.4.2.1 Effects of low intelligibility in communication the field

Low intelligibility is a particular issue for fieldwork communications in disaster management, as highlighted by TWB (2017a), when communication between groups was inhibited or inaccessible. Intelligibility is realised in a spectrum, ranging from total unintelligibility (when there is zero communication) to mutual intelligibility (when there are no communication issues) (Gooskens et al., 2015); in practice the aim is to have mutual intelligibility between all parties involved in disaster management, whether

that be the refugees or the aid workers. If there are no communication barriers, cooperation between all parties information can be achieved, allowing for the accurate transfer of information which can improve the refugee management system overall. However, if intelligibility decreases several communication issues can appear. For instance, when intelligibility is absolute zero, then information is simply blocked from being transferred, and as such paralinguistic gestures become the default primary method of communication (such as pointing the directions that a group should follow) (TWB, 2017a). The realisation of zero intelligibility is that even simple communication is inhibited, as there is no guarantee that the listener will be able to understand the intentions of the speaker (or pointer if paralinguistics are used).

The area of most research is the impact of low, but not non-existent, intelligibility; simply due to a range of communication difficulties being identified from the field in the situations. There are two key elements that policymakers need to account for when the intelligibility level, of the selected languages, is perceived as moderate (I.e., when people can for the most part understand), firstly, the lack of time available for sending messages and secondly the impact of misunderstanding in high stress high-risk environments. A common counter to intelligibility issues being perceived as serious, are techniques such as rephrasing, repeating or simplifying; speech (Yorkston et al., 1996; Rogers et al., 2010; Alves et al, 2020); these are seen as a way for communication difficulties to be worked around and adapted to. This counter isn't necessarily incorrect, as if there are no potential limits to time then intelligibility can be counteracted with effort, however in disaster communication by commodity

found sparingly. It is established across disaster studies, as a discipline, time is essential and that any delay to action decreases the effectiveness any attempt to reduce overall risk of harm to a population or individual (Carter, 2008; Vij, 2022). Therefore, if a message was required to be adapted so that the information can be understood, this requires time which, in the context of disasters, is non-existent. Policymakers must keep in mind the impact, that assuming that low intelligibility can be fixed, can have in the field, and in this case, it can lead to a lack of reduction to the risk of harm for the individuals involved, which is unacceptable from humanitarian perspective. The second factor, that of misunderstandings and the severity of miscommunication in emergencies must be accounted for in any language policy for disasters. In peacetime misunderstandings can result in confrontations or disruption, however in emergencies contexts the repercussions can directly impact on individuals chances of survival from a disaster event, in contrast to peacetime. Low intelligibility is when the nuances of specific information are not understood, consider evacuation direction orders (like tsunami warning alert) where the message itself is indicating both the severity of the situation (i.e., the risk currently facing population) and the direction that the population should travel in to reduce their risk of harm. Low intelligibility could result in critical words or elements of the message itself being missed or misinterpreted, which can result in an individual, or a group, going in the wrong direction than what was intended. In the context of active humanitarian disasters where the at-risk population are currently in the hot-zone, deviation from the intended route can result in a significant increased risk-to-life; especially in modern contexts when evacuation corridors are established through peaceful

negotiations (Holmes & de Piñeres, 2011; Doğan, 2016), and deviation from the corridors places the evacuees into the active war zone, where the risk to life is exceptionally high. As such, in practice, intelligibility can have real-world impact in disaster events.

2.3.4.3 How to measure intelligibility

At this stage, it is proposed to split the methods to measure intelligibility into two categories, based on the style of research relating to quantifying intelligibility . The first group is the structural approaches, which analyse intelligibility through pure linguistic differences within the languages explored, which compare the linguistic systems of languages to identify distance between all those involved. The second group is the performance approaches, which assess intelligibility through experimental methods, to record the responses to intelligibility related tasks or difficulties, allowing for the in-action processing of differences to be explored. Both approaches are beneficial for linguistic research overall, as objective results are recorded in consistent manners. Provided the research focuses on data-led approaches to language differences and intelligibility the goal of operationalising Abstand remains on course.

2.3.4.3.1 Structural

Lexical differences have been used as the criteria to measure the distances between language varieties, as well as exploring the processes that languages experience which cause deviation from the historic family language (Tang, 2009; Tang & van Heuven, 2015; Al-Jallad, 2019). Currently lexical differences are predominantly used

in language databases such as Ethnologue (Lewis & Simon, 2010) and World Language Atlas (Dryer & Haspelmath, 2013), whereby the differences are calculated by comparing the number of shared lexical items across languages. When assessing lexical differences, the approaches focus on the shared similarities as opposed to the differences. Examples of shared lexicon include the pronoun *ʔana* ('I' in English) across Modern Standard, Tunisian, Egyptian, Lebanese and Saudi Arabics, as well as *kta:b* ('book' in English) in Egyptian, Levantine and Saudi Arabics. The focus on similarities is logistic, as it is easier to identify shared lexicon using corpus data (see Bouamor et al., 2014; Khalifa et al., 2016; Kwaik et al., 2018b, Almujaiwel, 2020, Ibrahim, 2021), providing there is acceptable corpora to use, which cannot be assumed. Intelligibility measured through lexical similarities allows for investigations to focus on language change throughout time; for instance, consider European languages which are all derivatives stemming from the proto-Indo-European language, as such there are lexical similarities between modern European languages, such as English, Dutch and German containing multiple instances of shared lexicons (Gooskens, 2007; Gooskens et al, 2015; van Heuven et al., 2015; Gooskens, 2017; Gooskens & van Heuven, 2020). In practice, structural approaches to intelligibility can inform emergency language planning, as languages which share historical route or connections are often influenced by each other and overlap in part, with lexical items of the most shared factor in language contact.

Phonological differences can also be used as a measure of intelligibility between language varieties, particularly, the phonological realisations of shared speech features. An example of when phonological differences can be cross compared to

give an indication of linguistic distance, and intelligibility, can be found in Arabic, where the variation between Arabic varieties is predominantly found in the phonological realisation of Arabic (Munro, 1993; De Jong & Zawaydeh, 2002; Saiegh-Haddad, 2004; Trentman & Shiri, 2020). Modern Standard Arabic is the formal Arabic variety that replaced colloquial Arabic in usage, however the language itself is a taught language and thus lies in the phonological system of another variety of Arabic (usually the closest QA) for production (Albrini, 2011, 2014). If Arabic was to be used in the field, language planners could analyse the phonological similarities between different Arabic varieties, to assess whether a speaker of one Arabic variety would have a higher or lower intelligibility when heard by a series of different Arabic speakers/listeners. For instance, if there are significant overlaps between the phonological systems of the speaker and listener, then the likelihood that the listener will be able to understand the speech produced is higher than when the overlap is lesser (Gooskens & van Heuven, 2021).

2.3.4.3.2 Performance

Active speech ranking and replication are two areas of intelligibility research that can also quantify the linguistic distance between two or more language varieties (Keysar et al., 2000; Tang & van Heuven, 2009). This can include assessing whether an individual believes they can understand actual speech (which measures perceived intelligibility) (Tamburelli, 2021), the extent to which an individual can replicate heard speech (which measures both intelligibility in decoding speech and the intelligibility of producing speech) (Gooskens et al., 2015) and whether listeners can identify the speech community of the speaker (which measures the identifiability of a language

through intelligibility) (Munro & Derwing, 2011). Experimental studies exploring the intelligibility between speakers of two or more language varieties focus on the ability to decode and understand the messaging presented in the discourses when responding in the situation, or when making decisions based on topics in the messaging from the discourses. For instance, Gooskens et al., (2015) investigated the intelligibility of adult speech when heard by children, by having children listen to a range of basic lexical items in multiple language varieties, and after each word heard, asking the child to select image that shows the word, when the correct image was selected, it was scored as being representative of intelligibility within that word and when incorrectly selected it indicated no intelligibility with that specific word. Results of these lexical tasks were then compiled numerically to calculate the average intelligibility of the language varieties tested generally. The results of these styles of studies though do not necessarily indicate whether or not all speech is intelligible overall, rather it provides insight as to where intelligibility is greater and where it is lesser; allowing for research in the future to assess specific intelligibility scores in understanding based on controlled features, which were either intelligible in the child studies or unintelligible.

Furthermore, artificial distance has also been used as a method to calculate intelligibility. Tamburelli (2021) produced a study which exposed individuals to a form of English which had been artificially changed to deviate from standard English in a historical manner; with the aim to assess intelligibility when the differences are based on phonological processes that change the phonological realisation of a language through time, i.e., lenition and fortition of specific sounds in different phonological

environments. Through this technique, the intelligibility of a language was directly measurable, as the rate of distance between the forms of English tested were limited to being 5%, 10% or 15%. This was possible by changing individual sounds artificially using computer technology; and whilst these distances were arbitrary, they did provide an insight as to how intelligibility can be realised based on which aspect of languages the differences are realised in. Ultimately, Tamburelli (2021) found that as the distance increases the intelligibility decreases, however, the relationship between factors was not directly proportional or consistent across all speakers; in fact, on average there was a trend, but for the individual, the impact of linguistic differences varied. This highlights once more how intelligibility is not a clear-cut measure and calculating linguistic distance against intelligibility is even more complicated as it varies not only by which aspect of language variation is occurring in also on the speaker and their processing of the language.

2.3.4.4 Use of intelligibility

Throughout this section, there is one issue in intelligibility that has not been addressed. That of how close do language varieties need to be in terms of intelligibility to be considered the same language or mutually intelligible. Unfortunately, the literature is disputed on this topic and the exact thresholds of when to demark a language as a standalone language are unknown (Tamburelli, 2021). There are multiple arguments for multiple thresholds based on multiple features (Tosco & Tamburelli, 2021), which highlights the complexity both of language itself and of intelligibility, which is a measure of similarities between language varieties and the ability to understand across language varieties; issues which thus far have not

been unified across linguistics. However, the established intelligibility scores between language varieties can infer whether actual communication can be facilitated in an emergency, language policymakers can use the quantified results of intelligibility studies as a basis for expected levels of understanding between speakers. For instance, if the distance between two language groups is scored at 70%, then policymakers take the stance that at most 70% intelligibility is achievable through use of these languages in an emergency.

Additionally, speech intelligibility encompasses research conducted assessing the degree of understanding that can be achieved when the conditions are non-ideal for cohesive discourses, such as in noise (Ngo et al., 2020). In audiology, intelligibility has been used for testing and assessing hearing impairments (Clopper & Bradlow, 2008; Anand & Stepp, 2015). In linguistics, intelligibility in noise research has explored the relationship between dysfluency, background noise, interruptions, and language processing. Within noise research, several testing designs and systems have been developed for the testing of intelligibility for the hard-of-hearing and when speech is heard in noise, such as Kalikow et al's (1977) Speech Perception In Noise (SPIN) test, the Speech Intelligibility Index (Hornsby, 2004; Kates & Arehart, 2005) and the Malay Speech Intelligibility Test (MSIT) (Yusof et al., 2013). There has also been research on intelligibility and comprehension between natural human language and mechanical alerts (Hodoshima, 2018; Ngo 2020; McCloy, 2013). Intelligibility in variable background noise studies supports the understanding of emergency situations, which are by-and-large loud environments. Ngo (2020) identified that using mechanical sounds prior to the delivery of language stimulus improves the

intelligibility of the message overall, a feature attributed to the mechanical sound acting as a audio-reset, as it drowns out the other background noise. However, McCloy (2013) identified that natural discourses are experienced with a degree of background noise, and extrapolates that a lack of background noise can replicate the environmental silence following a loud mechanical noise, as people stop attempting to talk when there is a sudden and overriding mechanical noise experienced. A layman example is when a group of people talking amongst themselves in a small, enclosed environment can be silenced using a mechanical air horn; the messaging after this silence is the clearest to understand; whether this is due to the undivided attention given to the speaker, or a natural increase in intelligibility due to a lack of competing background noise is currently disputed, a topic which will not be explored further, as this thesis aims at exploring language use in inter-personal discourses, as opposed to group-based declarations.

2.3.4.5 Intelligibility in Arabic

Within the literature, it is widely established in Arabic studies that there are Arabic varieties which are not intelligible to others, such as Moroccan Arabic which is reportedly completely unintelligible to Levantine Arabic speakers (Bassiouney, 2014). The status of intelligibility within the Arabic speaking regions of North Africa and the peninsular is highly debated and disputed, predominantly due to a range of different techniques used to calculate intelligibility, namely the contrast between Abstand methods, (with linguistic differences: for phonological see Saiegh-Haddad (2003, 2018), Embarki et al., (2007, 2009), Farghaly & Shaalan, (2009) and Faeghaly (2010); for lexical see Saiegh-Haddad (2004), Zaiden & Callison-Burch, (2011, 2012,

2014); Zbib et al., (2012), Kwaik et al., (2018b) and Fiedler et al., (2019); for morphological see Gadalla, (2000b), Watson, (2007), Kwaik et al., (2018a) and Alruily, 2020)) and Ausbau methods (with socio-political identity: as summarised by Elmahdy et al, (2009), Hachimi, (2013), Sadat et al, (2014), Versteegh, 2014, Horesh & Cotter, (2018), Omar & Ilyas, (2018) and Al-Jallad, (2019)). It is established across the Arabic speaking continuum, intelligibility can be very low (Holes, 2003, 2006; Horesh & Cotter, 2018), however the intelligibility within the smaller gaps on this continuum scale are relatively under-researched. For instance, consider Mashriqi (or Eastern) Arabic, where there are multiple Arabic varieties in use across the region. There are multiple sources claiming different levels of intelligibility between varieties within Eastern Arabics, with some claims that there are little differences between the Arabic varieties in this group, and therefore claiming that there are no intelligibility issues, particularly when the Arabic varieties are used within the same nation and the same diglossic community (see Holes, (2003), Bassiouney (2009) Versteegh, (2014) for overview of socio-political views on Mashriqi unity). However, this is a sentiment not shared by Bassouiney (2014) who analysed the morphological similarities between three Mashriqi QA's (Levantine, Saudi (Gulf), Lebanese and Egyptian) against MSA and one Maghrebi QA (Tunisian), finding major deviations between both along the Mashriqi/ Maghrebi isogloss, but also between the Mashriqi's. Additionally, the Arabic unity sentiment was rejected by Zbid (2012) who identified lexical similarities through language use on twitter and advocated for the separation of colloquial Arabic's into distinct languages, a view reinforced by Alruily (2020)'s similar twitter analysis. As such, it cannot be assumed that all speakers within

Mashriqi Arabic groups will be able to understand each other. In fact, the research supports considering Mashriqi Arabics' intelligibility on a continuum itself, exclusive of global Arabic (i.e., the Arabic spoken in the Maghreb regions and nations outside of the Arabian Peninsula).

2.3.4.6 Conclusion

In summary, intelligibility can be used as the criteria to calculate the understandability of two or more languages in the field, however as it stands the techniques used to measure intelligibility and the impact of intelligibility in the field are relatively unknown. From the literature, it is established that the familiarity of a language to this listener is an important factor if high intelligibility is the aim, likewise the sense of urgency in the messaging itself is also connected to producing higher intelligibility rates across the population. However, the thresholds between marking a language as distinct or as part of another language are widely disputed and cannot be taken as fixed at present. Therefore, there is a research gap here, which warrants further investigation, as such part of this thesis will explore the following research question:

(2) Are the Mashriqi colloquial & standard Arabic varieties found within the diglossic continuum of the Kingdom of Saudi Arabia sufficiently intelligible, that it does not matter which variety is used in stressed contexts, such as in emergencies?

This sub-chapter has explored intelligibility, and the connection between the phenomenon and successful communication, both in general and in disasters. This review provided insight into the miscommunications and language barriers highlighted by TWB (2017a), namely, the accounts of interpreters being unable to

provide language support between two speakers, despite all speaking the same language, which is now established as a trait of Ausbaucentrism. Despite the advancements from this review, there is a remaining factor from TWB (2017a) which has not been explained, the status of attitudes within a community based on the identifiable features of the individual. In TWB (2017a) there were reports of refugees distrusting certain aid workers based on the language spoken. At this stage, what is unknown is whether this issue is unique to refugee populations or reflective of wider social attitudes, as such, the following sub-chapter shall explore the status of language attitude research, in relation to disaster language planning.

2.3.5 Language Attitudes

When delivering emergency communication and facilitating disaster management, it has been established thus far, in this literature review, that it is vital to consider the language choice and selection criteria, as well as the delivery systems and mechanisms for information en masse. However, there is a further element that has yet to be considered so far: the extent to which the information communicated will be believed by listeners, and/or whether the messaging will be complied with in the desired manner.

An approach of assuming that all messaging and information disseminated to an at-risk target population will be either received or accepted in full is ill-advised (Sandman, 1989; 1993). This forewarning is motivated from the literature on historic emergencies, disasters, and high-risk events, as it is not uncommon for socio-political rejection of information and aid to occur in disaster events (ibid). Historic

examples include the 9/11 New York 2001 terror attacks, when evacuation messaging was ignored (Sandman, 2012); the 1995 Rwanda Genocide, when peacekeeping orders were ignored (Fisanick 2004); and the 1995 Srebrenica Massacre, when the population denied the atrocity (Massey, 2021). The relationships between social trust, disaster management and aid management are well established (Sandman, 1989; 1993; Etkin, 2014; Pettit, 2015); particularly when there is a negative view between the community itself and the organisation represented. What has gained lesser attention is the role of linguistic attitudes in the complex mixture of sentiments, beliefs, and community stances that exist during active disasters.

This following sub-chapter will investigate the known socio-political relationships and linguistic attitudes between the Arab populations in the Mashriqi Arabic region. There will be four sections to this review. The first section will present outlines of current research and theory on the realisation of emergencies and the social responses. The second section will explore the status quo of Ausbaucentrism and Diglossia in Mashriqi Arabic social attitudes. Following this, will be a review into the use of language attitudes as a parameter to measure social trust between communities. Lastly, the final section will explore experimental methods to solicit underlying social attitudes based on linguistic criteria or features. The overarching aim of this review is to explore the relationships between Arabs shared social beliefs & the underlying language attitudes; with attention given to the realisation of such relationships when exposed to, or experiencing, emergency environments, contexts & communication strategies.

2.3.5.1 Social Grouping Theory

To understand the social responses of communities, and the individuals therein, to emergencies, it is important to first establish how said social groups are formed and maintained. This foundational overview is vital to understand the contextual factors, as communities are themselves affected by emergencies and disasters just as much as the individual.

An established theory to explain how societal groups and cultural norms are created, enforced, and encouraged is Social Identity Theory (SIT) (Tajfel, 1979). SIT (ibid) proposes that humans group and organise themselves according to social frameworks, with focus on distinguishing between those with shared connections against those without. If a group of people share a characteristic (e.g., language), physical identifiable feature (e.g., skin colour) or social identity (e.g., nationality), then they will categorise themselves as being part of one homogeneous group, which is labelled in SIT as an In-group. If a person interacts with members of this In-group but does not present evidence of shared features that makes the In-group, then this new person will be categorised as an outsider and placed into the SIT labelled Out-group. The relationship between individuals who are both part of the In-group is distinctively different to between In-group and Out-group members (Tajfel, 1979); whereas In-to-In relationships are more trusting, positive, and co-operative than In-to-Out, or Out-to-Out. It is important to acknowledge here that within a community there can be numerous social groups identified by individuals; and it is common for multiple groups to co-exist simultaneously within a single community group (Jacobsen & Landau, 2003). The stability and resilience of In-and-Out groups is questionable, as

these status's are dependent on the shared feature which justifies the group's formation, as well as the dominance of the shared feature as the primary cause for the social group itself (Tajfel & Turner, 2004).

Lower-level social groupings, such as community groups or friendship networks, are most exposed to changes of group dynamics and priorities (Abrams et al., 2005); with it possible for the dominant feature for the In-group to change with short notice; a process which is natural. For instance, the lower grouping could be formed by a set of office workers, who share a single office space, and that shared office space is the dominant In-factor. However, the introduction of an individual who works in another office space, but for the same employer, would establish another group, with the shared dominant In-feature being the employer. This example highlights how seamlessly social groups can be expanded or recalibrated without intentional effort being made.

Whereas, at the upper level of social groupings, the dominance of the shared factor is established and supported by the community (Hewstone & Giles, 1997); furthermore, there will likely be resistance to changing the dominant shared factor in favour of any other factor, with community members favouring the rejection of a new individual above their inclusion (Hogg, 2016). A common-place upper-level social group is nationality, specifically the shared feature of belonging to a nation state (Wright, 2016). The changing of nationality requires a formal process of documentation alteration, such as proof of citizenship, whereas, the regrouping of work colleague requires anecdotal proof of the shared work environment, such as a fellow colleague vouching for the new individual (Taijel & Turner, 2004).

There are positive and negative side-effects of socially grouping individuals; effects which are root causes for both social integration and segregation. Positive side-effects include acts of altruism towards helping fellow members of a community (Taijel, 2001) such as charity events. A further positive side-effect is the production and maintenance of strong inter-personal networks and friendships, which can improve an individual's mental health and attitudes towards the wider world (Skeldon, 2009). Inversely, the negative side-effects include racial segregation, where the communities are divided by a physical characteristic, extreme examples are found in colonial theories which argued for separating Rwanda into three classes, (Hutu, Tutsi, and Twa (Prunier, 2011)); ultimately setting up the ethnic divides that resulted in the 1995 Genocide (Besançon , 2005; Straus, 2006; Rawson, 2018). Additional negative side-effects include the more subtle and nuanced instances of discrimination (Tajfel, 2001), such as racist, sexist, or homophobic behaviour/language use, for example, rejecting the admission of a homosexual into an event. All three of these side-effects are a result of a new individual being considered different to the dominant In-group, and thus the members of the In-group utilise the differences as a vector to target and alienate or insult the other person; all to ensure that the outsider is aware that they do not belong within the community.

It is also important to recognise that a social attitude is made of a combination of judgements based on multiple distinct factors (Tajfel, 2004). A group attitude can be developed based on the shared characteristics within the group itself, this can be biologically determined factors (ibid), such as height or body size, it can be social grouped factors (Tajfel, 2001), such as being from the same employment sector, it

can also be based on linguistic factors, such as accent or language use (Holes, 2011); whereby the group identify specific realisations of sounds or conventions of speech as being indicative of being a member in the group (Bassiouney, 2009; Bidaoui, 2020).

On a greater scale, this can be applied to nationalities, ethnicities and political ideologies and identities, with the globe separated into political In and Out-groups, by virtue of nationhood and statehood. To create a nation, the population needs to believe that there is a shared sense of identity, in which sense all nations are large In-groups (Wright, 2016). Criticism of this level of the model, is that not all those who are in the national In-group category will perceive themselves as being the same, however in practice whenever this viewpoint is challenged, in a way where the nationality itself is at risk, social groups tend to unify by the national identity category as a single people (Hogg, 2016). For instance, this is part of nationalisation in situations of conflict and war, whereby a nations' population is more likely to become unified as one stable In-group if the nation itself, as the geopolitical entity is invaded and under threat of elimination (ibid). Populations band together to reject the invaders, who are perceived as total outsiders to the group, and in doing so form a stronger In-group for the invaded nation. This is an important factor to consider in human based disasters where conflict and civil disruption can be the mechanism through which the disaster event occurs or escalates from (see *2.2.1.1 Measure 1: Cause*), as in the sense of international disasters and conflicts the In and Out-groups of all the participating nations will be reinforced as a result of the conflict itself.

Within an In-group there can also be multiple other In-groups, such that a person may belong to multiple In-groups. In a simplified peacetime example, consider the notion of office politics, whereby all members are part of the same group by virtue of the employment they share, but within the environment they work, there are relationships and subgroups of In-groups that distinguish themselves from others within the same overall In-group. Such as there are friendship groups within the working group. On a greater scale this escalates to communities considering themselves distinguishable from their neighbours, whilst also both acknowledging that they are part of the regional society. The decision could be based on negative perceptions of others who share similar characteristics rather than on the individual in question, such as distrust of their ethnicity rather than a distrust of the person themselves (Mummendey, 2001). The decision could also be a result of bias, (Tajfel et al., 1979) where certain members of the community reject changing any characteristics of the overall group, to maintain a higher social status or sense of power within the group. Furthermore, prejudice is also an issue that is prevalent in social attitudes, with preconceived notions and ideas, and, in some instances, decisions being made about a person without ever interacting with them (Tajfel & Turner, 2004). Social groups, and the way in which social groups operate are the environment where racism, and other forms of discrimination can emanate from; and it is important to recognise that these distinctions between people will become heightened in an emergency or disaster, because of the wider instability in the region, which in turn heightens the emotional response for interactions (Sandman, 2012).

When a person is rejected from being part of an In-group, in that they are consistently perceived as being 'Out', the rejection itself can result in extreme consequences. In relation to language attitudes themselves, there are recorded instances of direct disaster occurring, in part, because of an individual being rejected from being 'In' by virtue of the way they spoke. For instance, there is the case of selective muteness with Virginia Tech mass murder in 2010 (Knoll, 2010) and the case of accent issues with New York State immigration centre massacre in 2009 (Bhatia, 2018). These extreme examples highlight how attitudes can develop into being deeply negative and encouraging or facilitating as a justification for performing violence sufficient to be classified as a disaster.

Furthermore, it is important to note that it is not just individuals who can be rejected by an In-group; organisations, corporations, and businesses can equally be rejected based on socially inspired judgements (Coombs, 2007). A community can decide to boycott the use of a business in response to the belief that the business itself is working against the group (La Macchia & Louis, 2016). The negative response can also impact societal relations between communities (Besançon, 2005), such as between In-groups and the government of the nation; these can be materialised by the In-group ignoring the instructions from the state, rejecting the information presented by actors of the state and sometimes be realised as civil disruption or disobedience with the intention of breaking away from the jurisdiction of the state (Reyntjens, 2009).

Anti-social negative responses are rooted in negative attitudes in society, with the communities holding distrust and a lack of shared In-Group values with the state

itself (Mir, 2019). In a disaster situation, where the state has a responsibility to reduce risk of harm to individuals in the danger zone, any pre-existing negative attitude towards the state or their actors, can still be realised and enacted upon by the communities (Coombs; 2007; Sandman, 2012). This can lead to the rejection of the veracity, or a refusal to follow any given instruction by virtue of the organisation that gives the order, or the individual themselves that represents the organisation holding characteristics that are perceived as being out of the In-group for the community but within the characteristic group of the Out-group of the organisation that is giving the orders (Massey, 2021).

If in a disaster there is a negative response from the community, specifically relating to the rejection of instructions or management techniques to mitigate risks to the population, then the severity of potential outcomes can vary. However, if the risk is related to short-term high impact consequences, then the negative reactions and rejection of disaster management instructions can result in injury or, in the extreme cases, fatality (Carter, 2008).

2.3.5.2 Social attitudes in the refugee populations

Social attitudes are also important to consider with refugee populations in aid camps or processing centres, as the negative attitudes can still be prevalent in the population, despite the changes in the localised environment, social framework, and political landscape (Jacobsen & Landau, 2003). Refugees in temporary accommodation are more vulnerable to extremism (MacVicar, 2020); for instance, if an individual presents themselves as an oracle of knowledge or wisdom, in that they can attempt to reassure the stressed, confused, and distressed individuals around

them; they can become highly influential in changing the opinions and beliefs of both the individuals and the social group overall. This has been realised in the Syrian aid camps, where Daesh supporting individuals have radicalised sections of the population inside of the temporary accommodation (Schulman, 2019), as such to promote a negative social attitude towards the international aid workers and the NGOs that they represent (ibid).

The status within the aid camps, is indicative of a schism between In-groups and Out-groups (Schulman, 2019). The Daesh supporters have created an In-group where the shared characteristic is the support of Daesh, and the sociological and religious ideological values that Daesh represent, in contrast to staff within the NGOs, who do not support Daesh and as such are viewed as an Out-group separate from the refugees. The repercussions of such a sociological framework is that there are multiple civic disruptions between the refugee groups and the aid workers (Clarke, 2018), including debates and disputes as to which organisation has jurisdiction for laws and control of regulations within the camp itself. Furthermore, there are also issues relating to communication between refugees in the camps and the aid workers, with refugees favouring to avoid telling the aid workers the truth of either the conditions in the camp (Schulman, 2019), groups acting within the camp, such as human smuggling (de Azevedo, 2020) or even to identify who is part of a Daesh supporting gang (Jones & Askew, 2016). This hinders NGOs mission to reduce the stress faced by the refugees, as the shelter that is provided, that are supposed to create a sense of respite and relief, has become tense and volatile.

An additional factor to consider in current aid camps is the demographics of the population inside the aid camps, given that refugee communities are composed of multiple social groups, nationalities, and ethnicities (Agustín & Jørgensen, 2018). It is important that when considering the refugees, the researchers do not consider them a single homogenous society, as in reality they are better represented as a cluster of groups which are unified by physical environment predominantly (Turner, 2016). When camps are considered this way, it can explain why there are underlying negative attitudes between social groups within the camps, which are based off the pre-existing attitudes that were held in the region of origin. For instance, a refugee from Iran may have negative beliefs towards a Saudi Arabian, who is also a refugee in the camp, based solely on the negative attitude of Iranians in Iran towards Saudi Arabians pre-disaster event (Mirshahidi, 2017); a view that has retained despite both individuals now inhabiting the same spaces and environments. This is a common place issue in refugee communities (Martínez, 2018), and is indicative of the magnitude that negative beliefs have on the attitudes of individuals, particular in that the beliefs are engrained far-enough to be retained despite social infrastructure that facilitated and encouraged said believes being either significantly damaged or no longer having any impact on the individual.

2.3.5.3 Status quo of Ausbaucentrism and Diglossia in Mashriqi Arabic social attitudes

Across the Arabic speaking regions of Northern Africa and Western Asia, there are multiple negative relationships between nations and their respective populations (Bakker & Singleton, 2016; Issaev, et al. 2022). These negative relations are

motivated by a range of different sociological factors, from religious differences to language production differences (Versteegh, 2014). Arabian speaking regions are Ausbaucentric and diglossic (Manfredi & Tosco, 2018), however, there are distinctions between other shared characteristics that are equally underpinning differences and negative attitudes. For instance, whilst most of the Arabic speaking nations are Islamic, there are denominational differences that mean that the worship of Islam is not uniform throughout the regions (Holes, 2011; Versteegh 2014); there are two primary denominations, that of Shia and Sunni, with the latter being the most prevalent across the overall Arabic-speaking region. There is a historical geopolitical distinction, between the eastern regions (Mashriqi) and the Western regions (Maghreb), which was formed from geographical distance, colonial influences and regional political alliances throughout the colonial era, and the post-colonial movements (Hachimi, 2013).

Linguistically, there are three primary distinctions within the Arabic speaking world (see *2.3.3 The sociolinguistics of Arabic: an overview*), each of which can be used to form social attitudes and prejudicial beliefs within the respective populations. There are the Mashriqi, the Maghreb and the International (Holes, 2011; Alansary & Nagi, 2014). There is the distinction between using colloquial Arabic and standard Arabic (Ferguson, 1991), with standard Arabic holding the H variety in the diglossia continuum whereas most colloquial Arabic's are L, and from that the use of colloquial Arabic in the wrong context can present a person as an outsider within their own community (Shouby, 1951; Al-Kahtany, 1997; Walher & Pederson, 2020). Furthermore, there is the distinction on a colloquial only level, in that there are

political rivalries for regional domination between two Mashriqi nations (Bassiouney, 2009, 2014; Samin, 2012) (Saudi Arabia and Egypt), and such the promotion of language varieties which are popularised in Saudi Arabia when around an Egyptian may be used as evidence, by the Egyptian, to exclude themselves from the same social group as the speaker.

Furthermore, there are historic and current conflicts which also add to the negative social status across the Mashriqi region. A noticeable example of current, is the Yemen civil war; whereby Saudi Arabia, a Sunni nation supports one side of the conflict, and Iran, a Shia nation supports the other (Bowden, 2019). This human disaster is being supported, in terms of weaponry, by both nations on each side, and as such the two nations are not friendly politically or internationally. This animosity between the nations is also reflected in the populations, with negative social attitudes being held within both population groups against each other; this is partly due to propaganda and nationalism but is also due to the conflict in which both nations are supporting opposing factions (Albirini & Benmamoun, 2022). Additionally, in recent history, there has been widespread social change movements across the Maghreb and Mashriqi regions via the Arab spring. The revolution resulted in constitutional changes for several Maghreb nations, such as Libya and Tunisia, however, the impact in the Arabian Peninsula was lesser so, such as in Saudi Arabia (Bowden, 2019). As such, the wider sociologistic landscapes were impacted; for instance, in Egypt the use of a local QA has become more acceptable than under previous governments, which has increased the status and power of the L varieties (Brustad, 2017; Hassan, 2021). This is notable as the same change is not reflectived in Saudi

Arabia, where the revolution was stopped (Bowden, 2019), and the state has resisted efforts towards social change. At present, the early work into the new sociolinguistic landscape suggests that the general public opinions held towards the state government are different based on whether the Arab Spring movement was successful (Hassan, 2019; Albirini & Benmamoun, 2022). In traditional states, the diglossia has remained stable, with the H variety being connected to courts and government, which are prestigious and highly trusted sectors, however, in the revolution states, the diglossia has shifted to improve the status of the L varieties, particularly the local QA's; and whilst the H variety is still used, the dominance of the variety is weaker than before the revolution (Brustad, 2017; Albirini & Benmamoun, 2022). As such, the diglossic status of a language (both H and L) may now vary by location; and early evidence is indicating that the social attitudes too are experiencing change.

Within the Arabic communities, the media language use can provide an insight into the stereotypes that are held across the region in relation to the language use of different groups. The primary Arabic variety used in Arabic-medium media is Egyptian (Frishkopf, 2010), particularly in the traditional fictional media, i.e., blockbuster films and television series; therefore, the Arabic speaking populations are likely to have been exposed to a range of attitudes which could be attributed to Egyptian speakers (ibid). For example, in a traditional action movie, there is the protagonist, who is good and virtuous and there is the antagonist, who is reprehensible and evil; when the language used is the same for both, the view could attribute other speakers of the language as being either positive (like the hero) or

negative (like the villain). Whilst this effect may sound simplistic, it has been established in the literature that the targeted use of a language variety in media can be a primary factor in attributing attitude onto a population (Hammond, 2007; Mastro, 2015, Al-Bazzaz & Ali, 2020); for instance, if a villain is exclusively using a single language, and the virtuous character speaks another language, the villainous characteristics will be attributed to the speech community that shares the same language as the villain. The impact of these influences can result extreme outcomes, such as overt racism (Dixon & Linz, 2002; Miller, 2020), however, in practise the majority of impacts are on the underlying prejudicial attitudes which are covertly realised, often without malicious intent (Mirshahidi, 2017).

It is important to recognised that within the Mashriqi region there are at least three Arabic varieties in general usage (Al-Suwaiyan, 2018). There are the traditional two; the local QA and the localised MSA, which exist in a diglossic relationship of L and H respectively (ibid). However, there are the alternative QA's used in the media, including social medias; of which the most prevalent is Egyptian (Kraidy & Khalil, 2018). The status of the media-led QA's is predominantly unmeasured or explored in research, mostly due to the speed of technological development of media; given that social media has become established in the past decade alone (Samin, 2012; Abdul-Mageed et al, 2018).

Within the diglossic continuum, the issue is where to place the media QA, and the classification is based on the level of familiarity and exposure. If the situation is such that the only exposure is exclusively via media channels, then the Arabic used may be perceived as either L or H depending on the content (genre) and context (formality

of the speech used) in the media programming itself. For instance, if the Arabic is only used by characters of low social status or within the domestic settings (i.e., at home, or with close family), then the Arabic would more likely be considered a L variety as the use reflects the typical context for a L Arabic variety (Stadlbauer, 2010). However, if the usage seen is from a high-ranking individual or group (in terms of social prestige, i.e., politicians, clergy, or judiciary) or within official context (formal news broadcasting), then the Arabic can be perceived as being a pseudo-H variety (Al-Kahtany, 1997; Versteegh, 2014; Al-Suwaiyan, 2018); as the variety cannot be the default H variety, as the use is fictional, and using the media-Arabic in the real-world official contexts is not accepted; therefore, the attitudes imply H status. As technology develops, the use of colloquial Arabics is growing, and as such, it is possible that there are various Arabics being exposed across multiple platforms and mediums to the modern Arab. However, the research into this field is limited at present, so to assume that a group will have exposure to an Arabic variety, through social media use, is to be advised against. As such, this thesis will proceed on the assumption of Egyptian Arabic as a primary Arabic in media, as there is evidence defending its status in terms of use in traditional media (Kraidy & Khalil, 2018).

2.3.5.4 Measuring language attitudes

To understand the political relations between Arabophone nations, current and historical, the attitudes of the communities need to be elicited and investigated. One approach vector for this is using language attitudes, to identify the beliefs held within the individual, and the compounded status quo within the overall community. There are two primary approaches used to identify language attitudes, there is the implicit,

where the underlying attitude is targeted for questions and experimentation (Bidaoui, 2020) and the explicit, whether direct attitude is targeted (Abbas et al., 2020). This thesis will consider the implicit methods for attitude solicitation, so that the target value, that of the underlying social beliefs and attitudes, will be identifiable and quantifiable; furthermore, the avoidance of explicit methods will avoid said issue such as bias (Guerini, 2008) and perceived belief versus actual belief (Schüppert et al, 2012).

Social attitude can be a useful measure to indicate the trust and belief systems and statuses within a population (Schiffman, 1998). A social attitude can be described as the thoughts towards specific individuals or sentiments held across a population group that shares similar characteristics in ethnicity, location, or social group (Schüppert et al., 2012; Heaton, 2018). There is a two-way distinction on how to collect attitude data and information, it can be either obtained from asking the group openly (Schüppert et al., 2016), and as a group, about their beliefs, such as through town halls, or through asking each member individually, exclusively of others, and then collating the information post collection (Derwing & Munro, 1997). In sociolinguistics, the latter group is preferred for research integrity as it allows for solicitation of honest feelings, emotions, and beliefs (McKenzie & Carrie, 2018); as individuals can experience negative emotions, such as anxiety (Frisch et al, 2015), when they are asked question in a group, when this occurs individuals are more likely to inform a researcher on what they perceive the overall community believes (Tajfel, 1979; Abbas et al., 2020). Whereas interacting with individuals singularly can solicit information as to what the individual believes, separated from what they think they

should believe. As such, this thesis will primarily explore methods and techniques to understand social attitudes which have been collected through individual testing rather than open group events, to ensure that the beliefs recorded are those held by the community rather than those they perceive they should believe.

Implicit techniques to solicit underlying language attitudes make use of deflection or deception in experimental design (Abbas et al., 2020). This is to ensure that the participant is not aware of the purpose of the experiment and testing, so that they do not produce responses that reflect the perceived beliefs, and rather allows for the individuals to be represented (Abdel-Rahman, 2016).

A study design that has been used across languages to explore language attitudes is guise studies, which is when a participant is exposed to stimulus without any additional information in relation to the source of the stimulus (Nejjari et al, 2019). These designs can be used with written (Mirshahidi, 2017), visual (Schuppert et al., 2013), or spoken (Schuppert et al., 2015) stimulus as well as for questionnaires (Abbas et al, 2020). Schoel et al., (2013), utilised Likert scales and guise testing to present the Attitude Towards Language Scale (AToL); a tool which was proposed to allow for the full evaluation of underlying language attitudes across three targeted dimensions, value, sound, and structure. The languages assessed across AToL studies were predominantly European, such as English and French, and as such is not directly comparable when assessing for Arabic varieties (Albury, 2020; Chakrani, 2020).

Another study that used a guise design is Mirshahidi (2017), who researched into the underlying language attitudes towards the speech produced by second language

Persian speakers from Iran. This study used spoken verbal guise as its design to solicit the underlying attitudes related to trust between native Persian speakers and non-native speakers. From this study the results indicated that there is a negative social attitude that exists within Persian speakers in Iran, when considering Arabic speakers. Mirshahidi (2017) indicated that Arabic speakers are perceived as the least educated out of a collection of five separate languages which are found in the wider area surrounding Iran; furthermore, the results indicated that there is a lack of positive sentiments in the underlying beliefs towards Arabic speakers. This further supports the notion that there is a negative status quo between different sociopolitical groups in the Arab regions. Abbas et al, (2020), used a match guise technique to explore the underlying language attitudes via the Language Attitudes Trait Scale (LATS), a technique that is comprised of 27 Likert questions designed to provide a full dimensional assessment of seven categories, highlighting how quantitative results can be compiled to explore the underlying attitudes. Abbas et al (ibid), also identified that languages when compared by In-groups do not necessarily always have to be distant, and in fact some varieties that were exposed can fall in the middle and be seen as a generalised midpoint, whereby the language itself is not overly positively nor overly negatively perceived, a neutral attitude zone can also exist in language attitudes.

Throughout language attitudes studies, ranking scales (such as Likerts and sliding scales) are an established mechanism for measuring and recording attitudes using quantified approaches and designs (Abbaad et al, 2020; Levon et al, 2021; Dong et al, 2022). These scales can be used to obtain data on attitudes discreetly, without

indicating to the participants what the specific testing objectives are, particularly when multiple factors are investigated, and the Likert questions are not directly related thematically. Examples include presenting a set of statements to an individual and requesting them to record their support or objection to the statements, and recording them on a scale, with support on one side and object on the other.

A further strength for the use of ranking scales in language attitude studies, is the ability to counter-act bias in responses, which can be obtained when the Likert themes are interspersed or reversed. Furthermore, ranking scales can provide a basic spectrum for analysis, whether that be through categorical intervals (such as Likerts) or through open-ended scales (such as 0-100), either approach allows for the attitude to be measured as a numeric value, which can be combined to indicate the average response to the selection between the binary options presented.

When focusing on the specific factor of the level of trust, or distrust, between the Arabic speaking populations of the Mashriqi there was little research identified. Bidaoui (2020) investigated the individual attitudes towards three Arabic varieties, Moroccan, Egyptian and Saudi using Likert scales, they identified that social categorisation is linked to linguistic cues in discourse markers. Furthermore, there results did identify a link between the social dimensions of solidarity with the linguistic behaviour and attitudes of Arabic speakers. This study focused on comparing Arabic varieties from both Maghreb and Mashriqi areas, so it can inform our understanding of the status quo of Arabic attitudes, and how language attitudes are represented in elicit responses.

At this stage, this thesis proposes that research be conducted to explore whether underlying language-based attitudes impact results, towards a general distrust of disaster instructions. This will be conducted to answer the following research question:

(3) Is underlying distrust between Arabic speaking communities a factor to account for in disaster communication planning?

2.4 Summary

Throughout this literature review, multiple aspects of emergency language policy, and general language policy, have been assessed and explored. The review has considered the status of international and national language policies, and language's selected in those policies, as well as the criteria for selecting a language diplomatically and regionally. The review has also covered the topics of intelligibility and language attitudes and explored the mechanisms for measuring the phenomenon and the impacts that personal and societal beliefs have on both. Furthermore, as the literature explored, the focus was retained on exploring how the situations faced within the refugee camps at the entry points to southern Europe have developed.

In sum, there are three critical elements of disaster language policy that are under-researched and in need of further investigation, which in combination can account for part of the issues faced in the aid camps, as a covariate. It is important to remember that no single factor can account for the complex situations in the refugee camps. The three issues are language selection, intelligibility & language variation, and

underlying language attitudes; each of the issues is reported as under-researched by TWB (2017a), and after conducting the literature review, this concurs with TWB's report.

Additionally, the three issues can be layered together, as they can happen in tandem or individually. When planning language policies for emergencies, there is the issue of language selection. Policies that favour the international diplomatic languages could be inviting a situation of disparity in language rights and access, specifically with the information being presented; in that the minority communities would be discriminated against or disadvantaged in comparison to the majority language speakers. Furthermore, there is no guarantee that using international language policies can provide language support for the impacted communities. If there is disparity in the language service provided, it can be argued that the human rights are being violated, and the UN presents language access as a fundamental freedom; therefore, an investigation into resolving this situation is warranted.

The second issue is language variation and intelligibility, as it is currently unknown whether a speaking population will be able to understand information presented to them when the language used is an Ausbau language. This problem is rooted in linguistic distance being an under-researched and undervalued element of communication. Furthermore, the impact of miscommunication in the field has been outlined, however, this disagreement across intelligibility research as to how to the criteria for calculating intelligibility and categorising mutual intelligibility. As there are acknowledged communication barriers in the field, which are a result of

Ausbaucentrism, it indicates that current language support is inadequate and therefore research is needed to improve the status quo.

The third issue is the impact of underlying language attitudes related to social unity and trust between speakers. At present, it is not established whether an intelligible message would be accepted as trustful Irregardless of whom delivers the message itself. In the field, there are accounts of social distrust between individuals based on judgements made on the speech practises or language used by an aid worker or interpreter. Thus, even if the language deployed is deployable, and intelligible, what is not known is whether it would be socially acceptable within the target community. This is a critical issue for disaster management; as failure to comply with evacuation orders can result in greater risk-of-harm to both an individual, and to those who follow in suit.

Therefore, this thesis will thesis will consider three research questions, in order of presentation, to provide insight and understanding of the issues themselves, whilst also presenting methodology to assess the extent of the issues, with the aim to support further effort to improve the status quo.

To recap the research questions are:

(1) Are international language policies suitable for use in emergency disaster language policies?

(2) Are the Mashriqi colloquial & standard Arabic varieties found within the diglossic continuum of the Kingdom of Saudi Arabia sufficiently intelligible, that it does not matter which variety is used in stressed contexts, such as in emergencies?

(3) Is underlying distrust between Arabic speaking communities a factor to account for in disaster communication planning?

Chapter 3: Assessing International Language Policy and Intervention Reach

This chapter will explore language policy and planning for emergency situations, specifically disaster events. The aims of the study will be to measure whether current international language policy is suitable for guiding the creation of top-down language policy for disasters. A new evaluation method for language policy will be proposed; *the Reach approach*, which is designed to explore the issue of language selection in policy, using objectivity as the primary approach. This is in-line with the ideology that in disaster response and management, the most efficient and objective approach should be taken as opposed to a bespoke individualised approach (Farazmand, 2001; Carter, 2008). In disaster management it is important to recognise the overwhelming risk faced by a population, and that risk is en masse, rather than focus on the individual (Carter, 2008). As such, this study will propose and explore the language policy in a case study design, measuring international language policies against a Reach based approach, to ascertain which method provides greatest language support to the impacted population by virtue of language selection.

3.1 Research Question

This chapter will aim to address the first research question proposed from the literature review: are international language policies suitable for use in emergency disaster language policies?

3.2 Aims

This study aims:

- To identify priority language choice for international intervention in situations of disaster, allowing for predictive language planning.

- To assess the extent to which language policy makers can use objective measurements to justify or guide language selection for international policies.

- To evaluate whether the current international language policies are acceptable for usage in current disasters, and whether the policies would have been acceptable in historic (and concluded) disasters.

3.3 Hypotheses

In this study there are three hypotheses, motivated from the state of the art.

1) Selecting languages for use in an emergency, via a Reach approach, will enable greater language support of the offset than arbitrarily using the languages of the UNLP; as a Reach approach will provide an element of bespoke-ness to the situation and will account for the regional languages at use in the disaster-impacted population.

2) A selection of the largest 5 languages in terms of Reach will not be sufficient in providing language support to the majority of the population, due to the prevalence of language diversity and minorities globally.

3) Assessing Reach will provide alternative choices for language policy formation than using current international and national policies as the primary guidelines, as an access approach will account for the deploy-ability of a language.

3.4 Evaluation Method

To assess or evaluate language policy within a specific context, what is needed is to establish rigor and a set of criteria for the selection and investigation of different cases and situations of policy, as it is important to maintain a consistent approach to allow for a comparable review of language policy. Within this study there are two parts of primary consideration which require attention regarding frameworks and criterion. The first consideration is linguistic methodology for assessing the effectiveness of language policy in the field, (akin to the phrase 'how to look' at language policy from a top-down perspective). The second consideration is defining the approach to case study selection, since the context in this study relate specifically to situations of human disaster, case studies will have to be measured and tested to ensure that any case that is selected is representative of human disasters containing war and conflict only (akin to the phrase 'where to look'). Within the field of disaster research, there are multiple types of human disaster, such as ecological and economic, thus a check-and-balance is needed to maintain assessing only the target context. Framework 1 will structure the first consideration, whereas framework 2 will do so for the second consideration.

3.4.1 Framework 1 – Case study analysis methodology

This framework establishes the criteria and the analysis of any selected cases within the review. There will be four stages of analysis within this framework. The first two stages will involve data analysis, the third stage will involve the collation of data and the final stage will provide summary commentary, based on the results of previous stages. Each stage will be defined and described below, for reference the order of stages will be presented consistently using the following titles:

A.   Measuring Reach

B.   Assessing Access

C.   Collation of Evidence

D.   Advising Future Issues

3.4.1.1 Measuring Reach

3.4.1.1.1 Defining Reach

In this stage, Reach is defined as being the total number of speakers within a target region who can understand a single language policy approach. In the context of this study, it would be when a disaster occurs, and the governing authority decided to only use a single language when interacting with the entire population of the region. If this approach is taken, then calculating the speaker density in the population will be able to present an indication of the population that will be connected with (or *Reach*-ed) in a target region. Ultimately using a single language approach is arbitrary in many nations, as multilingualism is the norm globally speaking (Guerini, 2008).

However, speaker numbers and speaker density can be used as a baseline measurement to identify the diversity within a target region and the populations therein. Using both speaker numbers and speaker density it is possible to identify the extent to which the community of speakers can be connected with, and utilized, from a bottom-up perspective but also a top-down approach.

Language density refers to the proportion of speakers in the total population, where it is common for multilingualism to skew the results. This is due to the number of speakers being counted as individuals; therefore, a bilingualism speaker is counted twice. This potential doubling of speakers can result in a speaker numbers equalling a greater number than the total population. In these cases, there is also another issue, the minority languages which are not counted.

Using density, when speaker numbers are greater than the population, could imply that there is a perfect dataset accounting for the whole population, however, this is simply due to skewing and the over-proportioning of bilinguals (by counting twice) may be eroding minority language communities from being viewed in the overall context.

To counter these issues, Reach takes the total population value as the sum of the total speaker numbers recorded, irrespective of the population in the region. This excludes the unknown values, and focuses the scope to including only the values which are known. By doing this, it allows policy makers the ability to plan the responses for an emergency based on the objectively accountable populations, with an acknowledge that there will likely be a lost section of speakers, but in emergencies, the planning must be focused on producing strategies for what is

known, and then the unknown can be handled in-the-moment. Whilst this isn't perfect, the overall aim of supporting the most with prepared response is still met.

Furthermore, by converting the densities into being respective of the sum of densities, the overall approach leans into the Abstand-model, with the objective values and known status's are used rather than the abstract. With Reach, the language demarcations will be explicitly based on Abstand-defined language varieties, where the intelligibility of the language varieties within the community is known, or at least considered as, acceptable for regular communication without additional support, i.e., interpreters. The use of density alone can result in ignored communities, whereas Reach aims to correspond the language varieties alongside established Abstand language classifications such as EGIDS (discussed later), allowing for the Reach to be measured numerically (as a proportioned density) and categorically with the EGIDS classifications, and both of these approaches allow Reach to indicate which language to use based on the size and infrastructure available for the deployment of a language policy.

3.4.1.1.2 Speakers as a resource

This approach was borne from Wright's (2016) notion that the volume of speakers is a core element in language policy in action. When planning language policy, particularly in disaster communication, resources available for each single language intermixed in a region should be accounted for. Speakers, as a resource, can be considered active and reactive. Speakers are not static and do adapt to situations (Ramirez et al, 2008), as opposed to pre-recorded sound messages for a telecom service which is static and presents continuous non-adaptive communication (TWB,

2017e). When considering that in disasters the primary workforce for communication is interpreters and interpretive services, one must always remember to account for the voluntary nature of such services. Furthermore, language policy and planning should also account for the reality that in-the-field communication is primarily the ad hoc discourse (TWB, 2017e), where the speaker acts as a de facto interpreter by virtue of knowing either one or two languages to a high proficiency. If the region affected is primarily monolingual, speakers also function as relays for information between governments, aid workers and the endangered population (O'Brien, 2016). Therefore, regardless of whether the situation is multilingual or monolingual, speakers can still provide a substantive boost to language policy and disaster communication by virtue of being facilitators for the relaying of information in the situation.

Speaker density can also be used to identify the extent to which pre-existing systems can influence the target population in a specific region. For instance, if you have a prebuilt sound system to alert the population of an incoming disaster and/or hazard (such as incoming weaponry), it is vital to consider how many people in the population will be able to receive messages. This is a critical problem for national level warning systems, language policy and planning (Lachlan & Spence, 2009; Bean et al, 2015). It is vital therefore for any policy to consider the intelligible reach of measures and messages within a hypothetical or existing region of disaster. A possible way to counteract the issue of speaker reception is to use the pre-existing speakers in the region as either interpreters (factor A) or relays (factor B), however, producing language policy reliant on both factors A and B, would be to accept that an

untrained volunteer workforce should provide or will provide sufficient and high-quality communication services to avoid harm. This seems arbitrary at best, as it does not provide any justifiable defence that the policy mitigates risk or accounts for any mitigated risk within the policy itself. As such it is vital to see the speakers as receivers of information as well as being an active resource to use as a de facto of being in situ.

3.4.1.1.3 Dataset Creation and Checking

This stage involved analysing language policy using two variables: 1) Speaker Numbers (SN) and 2) Population in the target region (PT). For each case study, a dataset of each variable will be collated. SN relates simply to the number of recorded speakers of a language. For this study, the SN dataset will use data extracted from Ethnologue (Eberhard et al., 2022), a database which contains demographic information on each known living language (7139 *et computatis*). PT data will be collated from the World Bank: World Development Indicators open-source database, specifically the indicator of Population [Total] (World Bank, 2019). The World Bank dataset contains demographic information via yearly interval entries, as such specific population counts can be extracted for each year. Whereas the Ethnologue (Eberhard et al., 2022) has fewer specific counts, as the dataset focuses on current and up-to-date counts of SN. As such, Ethnologue does not provide a mechanism to obtain SN for specific years.

To account for possible discrepancies, a recalculation of the dataset will be performed. If there is no recorded entry of speaker numbers for a target year, then the closest year entry (CSN) will be taken. The CSN will be then multiplied by the

population in the target region of a targeted year (PTY) over the population of the

year of the closest entry (CPT); thus, producing the data for the SN dataset for the

specific language. This will be used to calculate the SN of the target region in the

target year assuming that the demographics speaker numbers grew in line with the

overall population. For illustration, consider this in formulae notation below:

$$SN = CSN \times \frac{PTY}{CPT}$$

Furthermore, another issue to account for in the dataset creation is when the case

studies are international. The World Bank data repository stores on a country-by-

country basis (World Bank, 2019), as such, in multi-country cases, the population of

each country involved in the target region will be added together to create the PT.

3.4.1.1.4 Calculating Density

In this study, Reach will be measured using speaker density data. Density will be

used as an indicator of the saturation of a language in the overall population, which

can be used to assess whether the native resources in a target region are sufficiently

high enough to be used in a language policy for disasters. It is vital here to remember

the difficulties in language planning and policy planning when considering larger

populations, thus this indicator shall provide insight as to the start-point for disaster

policy creation.

Density will be calculated for each language in the target region. This calculation will

be fractionalizing the two variables (SN & PT), followed by a multiplication of 100;

producing a final value (a percentage) for density that will be between 0-100. In the

event the final value is >100, the data & data sources will be reassessed; given that a

score of >100 would indicate that more than 100% of the population in the target

region speak the language in question. For illustration, consider the formulae notation

below indicating the specific language):

$$(Density\ of\ \emptyset)_a = \left(\frac{SN_a}{PT}\right) \times 100$$

Furthermore, at this stage, it is important to consider two notions. Firstly, that the

world is predominantly multilingual (McMenamin & van der Walt, 2018), thus it is

likely that there will be individuals within the PT that will be accounted for in more

than one SN. Therefore, there is little concern for $\sum Density_n \neq 100$ (here $\emptyset\_n$

indicates all languages), rather it is expected that $\sum Density_n > 100$.

Secondly, the relationship between Reach and PT can indicate errors or issues in the

dataset. A situation with $\sum Density_n < 100$ indicates that the dataset is incomplete for

the language demographics of the PT. In the event that this occurs, the data, and

extraction sources, will be reassessed; if no additional data is obtained then the study

will continue with the density calculated thus far, albeit with caveats acknowledging

that the data is as complete as can be. Whilst it is unfortunate for the dataset's to be

incomplete, it is not unprecedented as the collection and recording of demographic

information varies greatly nation-to-nation (Ethnologue, 2022), particularly when

comparing developed nations to developing nations, as the latter populations are

more inaccessible due to a lack of technological infrastructure.

To provide a structured illustration of the status quo of language density, all case

studies will contain a tabulation of density scores. Where possible, the tables will

include all languages in the target region. In the event of high multilingualism, the

tables will be altered to show the 10 languages with the highest density, with a 11th category for the collated density of all other languages in the target region. For an example, see table 3.

Table 3: Density scores for an example case study of high multilingualism

Language	Density (%)
A	51
B	43
C	21
D	12
E	8
F	6
G	3
H	2
I	2
J	1
Other	17
Total	166

3.4.1.1.5 Calculating Reach

A value for Reach will be calculated for each recorded language in the target region. The Reach of a language will be calculated by using the density of the language in question, against the total density of all languages within the target region. For illustration, consider the calculation in formulae notation:

$$Reach_a = \left(\frac{Density_a}{\sum Density_n}\right) \times 100$$

This will be conducted to produce a prediction chance for each language in the target region. For clarification, the prediction chance relates to if a single person were to be randomly chosen from the PT, what is the likelihood that the selected individual is a speaker of any of the languages in the target region. This prediction does not address whether the selected person is a multilingual speaker since the approach is designed to address the question of whether a randomly selected individual would be able to understand communication in a specific language. Converting the density into Reach allows for a clearer picture of the status quo of the language and can be used to identify how effective a language policy could be at a maximum.

It is important to note here that if density alone were to be used, in replacement of Reach, there would be an underlying issue relating specifically to when $\sum Density_n \neq 100$.

Density, as a single measure, cannot provide inference to address the question Reach addressed. Using the example data from table 4, the level of multilingualism can be suggested to be as high as 66, which suggests that 66% of the population are bilingual. However, the data (aka the Density as a stand-alone measure) cannot say

either way as to which languages contain the bilingual speakers. Whereas, Reach adjusts the density to account for this issue; converting the raw density into a comparable percentage chance (given the restriction that all values must fall between 0 & 100). For illustration, consider the Reach scores in table 4 alongside the raw Density scores in Table 3.

Ultimately, when considering a language to use, the aim within this section of the framework is to identify the largest language, in terms of population, within the PT. Both Density and Reach will produce the same order of language to use, given that Reach is an adaption of Density. However, the additional step of converting density into Reach, is conducted to produce more accurate detail for consideration in the later steps of language planning.

Table 4: Reach calculated for an example multilingual case study

Language	Reach
A	30.72
B	25.90
C	12.65
D	7.23
E	4.82
F	3.61
G	1.81
H	1.20
I	1.20
J	0.60
Other	10.24
Total	100

3.4.1.2 Language Access

3.4.1.2.1 Defining Access

In this study, Access is defined as the availability of identifiable, understandable, and trustworthy information that can be produced and complied within an emergency by a target population. Measuring Access will allow for insights into the preparedness of a

language policy for sudden emergencies. Furthermore, quantifying Access will allow for objective comparisons between language policies & the regions in which they cover. Access is measured to understand and calculate the extent to which a language can be deployed to a target region. The deployment of a language in an emergency involves producing and releasing information in a target language to a target population in a short-time window (Gardner-Chloros et al., 2016). It is important to note here that the context of information will be based on a fluidic and newly emerged situation that requires immediate action from the population.

For this study, a set of factors to consider in language policy making will be presented. Each presented factor will explore the limitations and issues within each said factor, and the impact it can have on language planning and use in disasters. Furthermore, methods to record or account for each factor will be presented in the accompanying sections. Where possible, a pre-existing academic categorisation system for language will be identified to use to calculate each factor, with the intention of using a pre-existing database. This approach is borne out of the difficulty facing language policymakers, particularly those looking at international levels, as the collation, collection & recording of all the language factors and demographic information takes intensive time and resources, thus it is acceptable to use pre-existing databases rather than start anew. Upon each factor being considered, and a solution for calculating presented, the final product will be a method to evaluate Access for language policies in disasters.

3.4.1.2.2 Factor 1: Linguistic Diversity

Firstly, there is the notion of producing resources for every possible language, with this the primary issue is the linguistic diversity. Furthermore, whilst it would be ideal to produce language resources in advance of every language in a target region, logistically it is unfeasible. The majority of the world's languages are minority languages, which means they are spoken by a smaller (and sometimes endangered) population, as such accounting for small groups of people is a consistent issue faced for language policy makers (Johnson, 2013), particularly when accounting for the contexts of disasters (Bolin, 2007; de Varennes, 2021). Globally there is widespread disparity between the majority languages and the minority languages in terms of resource stockpiles available for the replication, production or education of a language (Cenoz & Gorter, 2017). Majority languages, by virtue of having more speakers, have a larger stockpile of education textbooks, standardised dictionaries & printing ability for written text, than minority languages (Wright, 2016). Whilst this is something linguists strive to one-day eliminate (Kazeem & Suleiman, 2020), current policy makers and research must remember to consider the disparities as a matter of fact. Therefore, it is vital to identify the majority languages from the minority languages when assessing language policy in disasters.

Languages can be categorised as majority or minority based on a range of factors, ranging from size of speaking population to level of inter-generationality (Loos, 2007). In this study, the vitality of a language will be considered, going forward, as the start point for any classification of variation. Vitality is chosen primarily due to the nature of the disasters, as initial planning (for prospective future disasters) needs to assess

whether the language will still be in the target region & will still retain the same levels of existence. This is not to say that only languages which cannot change status, stockpiling & population can be selected for disaster language policies, rather the focus is on accounting for potential changes in a language in the future, as well as the status of the language at present.

When selecting a vitality scale to use in this study, the start point will be assessing the United Nations Educational, Scientific and Cultural Organization (UNESCO) (2003) language vitality and endangerment survey for suitability. The UNESCO (2003) vitality scale focused on identifying and categorising languages as either a majority or minority language as a start point, before further categorising minority languages to different levels of risk, whereby risk equals the chance of going extinct. A combination of six factors were used to assess the vitality, furthermore, highlighting how multifactor approaches are standard in language policy. The factors chosen included material for literacy & intergenerational language transmission. The UNESCO (2003) scale provided a six-scale classification (0-5), with a focus of the scale being to highlight low levels vitality rather than high levels. For instance, when classifying the first factor (Intergenerational Language Transmission) (ibid: pp.7) of a language, the terms used to correlate with the grade given (0-5) are positive for only a score of 5, with the term 'safe'; whereas grades 1-3 each contain the word 'endangered' with differing adjectives which increase in severity as the scale decreases. The final score of 0 is reserved, across the UNESCO factors, to classify languages which are lost or 'extinct' (ibid: pp.10).

The major criticism of the UNESCO vitality survey is that there is particular bias towards languages in decline (Lewis & Simons, 2010). It is worth remembering, that the UNESCO vitality survey was the start point for fulfilling the task of classifying and identifying languages and scaling them objectively using a sociolinguistic framework. The scope of the UNESCO (2003) report originally was to provide a framework from which to develop language support systems from, rather than provide comparisons on language strengths. The UNESCO vitality survey built upon Fishman (1991) Graded Intergenerational Disruption Scale (GIDS), which focused on classifying the transmission of a language from one generation to another, therein named inter-generationalism; since Fishman (1991) highlighted that inter-generation transmission of a language was a vital factor for a language to survive. The GIDS (ibid) is an 8 point-scale, with the highest-ranking languages described as so due to usage in education, work, government on a national level, whereas the lowest scores were applied when the language held a declining speaker population, with many of the speakers being of the older generations. Criticism of Fishman's GIDS is like that of the UNESCO vitality survey, specifically in reverse as Fishman's GIDS focuses too far on the strength of the language (Kaplan & Baldauf, 1997), or as a positive bias towards oral languages.

Lewis & Simon (2010) produced an updated classification system for language vitality focusing on utilising the positives of both the UNESCO survey and Fishman's original GIDS. This updated application system is a 11-point scale entitled the Expanded Graded Intergenerational Disruption Scale (EGIDS). The EGIDS was an extension of Fishman's 8-point scale, with the aim of improving the scaling for

multiple perspectives. The 11-point scale was designed for assessing language loss (as the approach to calculating vitality), however, there is an alternative ranking when considering language revitalization; whereby six of the scale points are converted into different category terms with different descriptions (Lewis & Simon, 2010). For example, when considering language with quality from a loss perspective the EGIDS score of 8a is considered Moribund, which is when the language is only spoken by the elder generation, with a decreasing speaker population occurring; however, when considered EGIDS from a revitalization perspective, the score 8a becomes Reawakened; with the acquisition of the language in the community spaces, with oral use of the language in the context of domestic usage. Whilst the perspective differences are not directly relevant to this study, as a method of calculating vitality, it is an indication of the objectivity and usability of EGIDS in language policy in language planning.

When assessing a language using EGIDS, 5 key questions are considered (see table 5). The questions which can be answered, and the answers given therein are used to identify where on the EDIGS a language can be categorised at. For instance, for the lowest two levels of the EGIDS (9 - Dormant & 10 - Extinct), only the first question (see question 1 in table 5) is answerable, with the level of answer selected used to decide between the two levels (9 & 10). This decision-making process is regularised and objective, thus the EDIGS will be used as the primary measurement for assessing access.

Table 5: Questions and Order of Questioning for EGIDS Evaluation

Question Number (with lowest asked first)	Question
1	What is the current identity function of the language?
2	What is the level of official use?
3	Are all parents transmitting the language to their children?
4	What is the literacy status?
5	What is the youngest generation of proficient speakers?

Furthermore, Lewis & Simon (2010) mapped the EGIDS against the UNESCO; providing a comparison. This mapping highlights how EGIDS is an expansion of both Fishman's scale and the UNESCO's scale (see table 6 for comparison between the three scales). Given the scope of EGIDS, to use the scale to assess access, a reduction of levels which can be used is warranted. Initially, the languages with an EGIDS score of 0 will be removed, since the score relates to languages used on a global perspective and level; whereas this study's remit limits to specific target regions; thus, EGIDS of 0 is incompatible. A further reduction of EGIDS levels for use will be conducted using the remaining factors (2 & 3) within the framework as presented above.

Table 6: Comparison table of language vitality scales.

Theta (Ø) indicates the comparison is unavailable.

GIDS	UNESCO		EGIDS	
Level	Level	Label	Level	Label
Ø	5	Safe	0	International
1	5	Safe	1	National
2	5	Safe	2	Regional
3	5	Safe	3	Trade
4	5	Safe	4	Educational
5	5	Safe	5	Written
6	5	Safe	6a	Vigorous
6	4	Vulnerable	6b	Threatened
7	3	Definitely endangered	7	Shifting
8	2	Severely endangered	8a	Moribund
8	1	Critically endangered	8b	Nearly Extinct
Ø	0	Extinct	9	Dormant
Ø	0	Extinct	10	Extinct

3.4.1.2.3 Factor 2: Language Stockpile

Disaster language policy needs to account for the current stockpile level of resources, and to account for the potential level of resources that can be operationalized or created in the short term (Edgington, 2021). However, language policy makers must also remain aware of the fluidic nature of disasters, where the information required for the general population, such as evacuation orders, can change on an ad hoc basis with very little notice; consider Tornado warning systems whereby a tornado can change direction rapidly (Britton, 2007). As such, language planning and disaster planning can create some resources for use such as generic evacuation orders, but the specific and targeted interventions cannot be prepared for in advance. The other consideration is that for a language to be deployed there needs to be sufficient infrastructure to produce the ad hoc message. Additionally, language infrastructure great enough to provide information to a large body of people in a very short amount of time requires investment and maintenance consistently, which is economically difficult to justify for every single language in a target region (O'Brien, 2017) (particularly if majority of the languages are minority languages). As such, what needs to be identified in each target region is the language which has the greatest potential, and current capacity, to produce language resources en masse at short notice.

When applying these considerations for calculating language stockpile, the scope of EGIDS in this study is reduced from allowing most language classifications (i.e.: 1-11 scores), to only allowing languages scoring between 1 – National and 4 – Educational. This limit was selected as the criteria for categorising as languages

classed at level 4 when language is use as the medium for the public education system as such, most of the population will be able to understand the language in a written format.

3.4.1.2.4 Factor 3: Language Status

The primary consideration of language status relates to the context in which a language can be used in the target region. This includes whether the language can be used in court, government & trade. The context of use for a language has been found to impact mobility and trustworthiness of information (Ng & Hamby, 1997) particularly in emergencies (Ramirez et al., 2008), as well as being an influential factor for the compliance with messaging (Taibi & Ozolins, 2016). For instance, if language is used in trade, and the speakers in the community often hear the language during trade, they are more likely to consider the speaker part of the social In-group, and thus trust them more so than if the speaker was identified as part of a social Out-group (Tajfel, 1979). Therefore, it is important to account for the social acceptability of a language in the target region, whereby the use of the language does not appear to be a faux pas in the community.

When considering the EGIDS scale, a further restriction in place will limit the selection of languages to only those ranked between levels 1 – National to 3 – Trade. Given that a level 3 (or lower) classifications requires a language to be usable in a workplace without consequence, and therefore the language community are considered part of the In-group; thus, accounting for social In and Out-groups within the scale itself.

3.4.1.3 Collation of Evidence

To assess the language policy in a target region, this study will present a list of the top five languages to deploy in the target region, when considering an objective perspective. This list will only include languages which achieve the Access criteria (i.e.: EGIDS 1-3); and the order of languages in the list will be based on the Reach score, in descending order. Given the importance of considering the feasibility of a language policy, particularly in a disaster, only the top 5 languages will be presented in the final list. It is important to remember that the top five languages will most likely not account for 100 percent of the population in the target region, this is accepted as a constraint of this approach. However, it is justified as it is a start point to build more adaptive and predictive disaster language policies.

The final top 5 list will then be assessed against two core criteria, which will evaluate the potential of a language policy which uses the top 5 languages. The first criteria relates to the preparations and development potential of a language policy by comparing the Top 5 to the use of the UNLP in the same target population. If there are any UNLP top 5, it can indicate a level of preparation internationally, which could improve the production of pre-recorded materials if the nation used the UN's assistance. Furthermore, a UNLP language would be easier to learn for the population en masse, primarily due to the framework for education already being in existence for the 6 UNLP languages. These education infrastructures can be utilized to teach the languages faster than smaller, more localised languages, which would require the teaching resources being produced and then being used, an additional step when compared to the UNLP languages. For the avoidance of doubt, it is worth

noting that the 6 UNLP's languages are not easier to learn for an individual than the other 7000 languages on the planet.

The second criteria relates to the acceptability of the Top 5 languages, including whether the top 5 Reach's when combined amount to significant coverage of a target population (with significant set at simple majority of 50%+), as well as, whether there are disparities or noticeable gaps between the Reach's of the languages in the Top 5. This criteria addresses is disparity between the usability of the majority languages in the target region, to illustrate consider this example situation; whereby there are 3 large majority languages in the top 5 (with Reach scores of over 20%), but the final two languages in the same list account for less than 5% of the population. In this case, there is a logic to using all 5 languages, as they are all the largest, but from an economic and logistical perspective, there is also the argument to use a tri-lingual policy first, which would account for over half of the population. The feasibility of a language policy in theory is difficult to quantify, with this study, the goal of for policy is if the languages selected could support a minimum of 50%+ Reach by using the least number of languages. Whilst this target appears arbitrary and failing to reach the UN goal of providing language support for all, it is important to remember that currently language support is 1/13, as such, an improvement goal of raising support to ½ is a pragmatic start-point for language planners.

3.4.1.4 Advising Future Issues

The lists provided using this methodology, can inform policymakers when developing future disaster policy and for evaluating existing policies. For instance, if only one language is identified as acceptable to use via this method, but the language does

not account for the whole population of a target region, it would signal to

policymakers that expansion of other language frameworks and infrastructure would

be required to access other parts of the population in a systematic manner.

Moreover, if the top five languages are selected but there are more languages that

could satisfy the Reach criteria, it would indicate to policymakers of the extent to

which diversity in a linguistic sense needs to be accounted for in the pre-emptive

disaster planning. This can be achieved by commissioning more pre-made resources

for the languages which do not feature in the top five.

Ultimately this method aims to provide an objective baseline for further analysis of

each case study this method is applied to, it is not an exhaustive measure in itself,

rather, it is a justification for further assessment in preventative disaster policy and

planning.

3.4.2 Framework 2 – Selecting Case Studies

Framework 2 will outline the criteria for selection of a set of case studies to apply and

evaluate the methodology presented in the previous framework. Each of the case

studies will explore the current and potential language policies within a disaster

setting. There will be five criteria presented, namely: (A) Type of Disaster, (B) Scale

of Event; (C) International Aid Intervention, (D) Multilingualism, and (E) Data

Availability. Additionally, for a well-rounded comparison, at least one case study will

be required for each of the following temporal instances: Past or Present.

Furthermore, a case study will be selected to assess preparation level of a nation, in

the event a disaster occurs in the future; it is vital to note that this report will not state

in any way that a disaster is, or should be, occurring in the future to the selected nation.

3.4.2.1 Type of Disaster

There are multiple types of disasters, which are firstly dichotomised into two core groups: Human-based and Natural (see *2.2.1 Defining and categorising human disasters*). This study will only consider human-based disasters, to limit the variation between case studies; since it is difficult to compare between disaster core groups. Furthermore, this study will also limit the case studies taken to human-made disasters, whereby the primary risk is created by a human-led action or actor. Examples of human based disasters include Nuclear Fallout, Genocide & Civil War. In this study, a further distinction will be followed; in that, only disasters with a direct and active actioning by humans with the intention to perform warfare or human harm on a specific targeted cross-section of the population; and as a result, the risk is directly related to the actions of humans (such as gunfire or artillery) will be included. From this distinction, only two types of human disaster will be considered, firstly will be situations of genocide, whereby the actor operates with the intent to destroy a demographic group (such as ethnicity or religious), and secondly, situations of civil war, whereby the warring factions, which populate the same region, are operating with the intention of injuring the opposing faction.

3.4.2.2 Scale of Event

In this study, the scale of disasters to analyse will be limited; whereby scale refers to the size of the disaster in relation to the geo-political landscape. The lower end of the

disasters to be assessed (smallest event) will be national-level disasters; by this, the disaster must have an impact on an entire country or sovereign region. Whereas the upper end will be international but mono continental; whereby, the event is limited to a single continent. This is to regulate the comparisons; given that the term disaster can describe a wide range of events, which can occur from locally to globally. For clarification, a nation-state in this study will be defined using the United Nations list of recognized sovereign states (UN, 2022).

3.4.2.3 International Aid Intervention

This study aims at assessing major disasters and disaster events. To facilitate this, the case studies taken for examination will be screened for their international significance. In the cases of disasters of the past, and with those of present day, this criterion will be satisfied if there were/are an international aid effort to provide one of the following: Providing peacekeeping forces and negotiators for warring factions, and/or supply aid-delivery to support the displaced populations.

3.4.2.4. Multilingualism

Any case study taken must contain a population that speaks more than two languages. There will be no ceiling on the level of multilingualism applied, primarily due to a lack of rigor justifying a ceiling (as any ceiling applied here would be arbitrary).

3.4.2.5 Data Availability

The final check and balance for any case study going forward is that there is an available dataset for the targeted region of the disaster, specifically that there is

language demographic data that can be either taken at face value or be converted to be representative of a different time. Whilst this goal seems simplistic, it is important to recognise the difficulty of obtaining language data for significantly minority communities i.e., those in remote regions disconnected from the main population or government.

3.5 Case Studies Selected

In this study there will be four case studies analysed, with the purposes of assessing and to test the proposed methodology in this study (*see 3.4.1 Framework 1*). Each of the four case studies satisfy the outlined criteria above. There will be two case studies taken from historical disasters, one from a current disaster and one case study shall act as a predictive target region possible disaster. Furthermore, three case studies will assess single nation disasters, with one case study therein for each temporal criteria (Past/Present/Future) and there will be one case study for multiple nation disasters. When combined the four case studies will assess the language policy of use/ or potential usage, using the outline method; this will allow a two-pronged analysis, first of the language policies and second of the method presented previously (*see 3.4.1 Framework 1*) (see table 7 for list). The four case studies will be the 1994 Rwandan Genocide, the 1998-2002 Second Congo War, the 2011- Syrian Civil War, and the Nation of Ghana (not currently in conflict).

Table 7: Case studies selected with criteria for selection satisfied & matched-up.

Case Study	Type of disaster	Scale	Intervention	Multilingual	Data	Temporal State
Rwanda 1994	Genocide	National	Yes	Yes	Yes	Past
The Congo 1998 – 2002	War	International	Yes	Yes	Yes	Past
Syria 2011 -	War	National	Yes	Yes	Yes	Present
Ghana Future	War	National	Yes	Yes	Yes	Future

3.5.1 Case Study Analysis

The following sections contain the analyses of the four case studies in turn. The case studies are ordered into the following sections: Background of the event, Results, Comparisons and Summary. This order was designed to provide regularity to the comparisons, whilst allowing for comparisons between the studies to be made more consistently.

The results section will outline the Reach for the event. The comparisons sections outline the differences between the top 5 languages by Reach, against the UNLP and

the national language policy enforced prior to the disaster event. The summary will outline trends in the case study.

Following the four case studies, a summary analysis was conducted, with the general trends of weakness and strengths of the UNLP approaches, national language policy approaches, and Reach approach calculated and investigated. The summary also identified the weaknesses of language policy from a top-down perspective, and whether any of the approaches achieve the basic human rights required for an equal or equitable language situation.

3.5.1.1 Case Study One

3.5.1.1.1 Background Information

The Rwandan Genocide began in April 1994 and concluded in July 1994. The disaster directly began with an assassination of the Rwandan president Habyarimana on April 6th 1994. Whist the genocidal event was limited to 1994, there event was rooted in political instability and discontent with the diplomatic solution and peace agreements concluding the Rwandan civil war, namely the Arusha Accords (Rawson, 2018). The violence was primarily aimed at those of Tutsi ethnicity (Magnarella, 2005), an ethnic group which enjoyed social and economic prosperity during the German and Belgian colonial eras (Straus, 2006); and had retained that status, holding political power despite being less than 20 per cent of the Rwandan populations (Miller et al., 2020). This social inequality was used to justify the slaughtering, raping, and torturing of the Tutsi population by the pro-Hutu militias (Hintjens, 1999: Magnarella, 2005). Following the assassination there were violent

militia-led campaigns of violence, massacres, and ethnic cleansings (Rawson, 2018) the primary goal of the militia. The fatalities from the conflict are estimated to be at least 800,000 Tutsi and 10,000 Hutu, around 11 per cent of Rwanda's population at the time (Magnarella, 2005: 816). By the conclusion of the genocide, an estimated two million people were displaced from Rwanda into the neighbouring countries (Straus, 2006).

Prior to the genocide, there were three recorded languages: Kinyarwanda, French and English. The linguistic landscape, and socio-linguistic dynamics were complex. French was the language delivered to the nation by the Belgium colonialists; English was introduced through trade with other African colonies and Kinyarwanda was the native Bantu language. The colonial languages held higher prestige in state-matters than the native Kinyarwanda, this was evidenced by the Arusha Accords being produced in only the colonial languages.

During the genocide, language was used to incite hatred towards the Tutsi population. This was evidenced through the radio communications leading up to, and during the genocide (Straus, 2006). There were more French speaking Tutsi's than Hutus, and the ability to speak French was, in some parts, sufficient of an identifier for targeting (Li, 2004).

3.5.1.1.2 Results

For this case study, the target population required was that of Rwanda pre genocide. The closest full record in the World Bank's repository for Rwanda pre-genocide was 1993. There were three languages reported as in general circulation at that time

period. The Reach was calculated for the country circa 1993 (see table 8). As the total number of languages was lesser than five, the condensing to the top-5 step was not required.

Table 8: Language Reach for the Rwandan Genocide (all rank 1-3)

Language	Reach
Kinyarwanda	91.62
French	8.12
English	0.26

3.5.1.1.3 Analysis

3.5.1.1.3.1 UNLP

There were two of the languages of the UNLP present in the nation, both of which are the working languages. Despite this, if the UNLP was applied and used as the guidance for any peace settlement efforts, or as the language in refugee management, then less than nine per cent of the population would be catered for. However, if the native language was employed singularly, then the Reach would have been staggering at over ninety per cent.

3.5.1.1.3.2 National policy

Given that the de jure languages were English and French, the national policy would have been insufficient, exactly like a UNLP approach. However, given the dominance of Kinyarwanda, it is highly doubtful that any other language would be used. The

issue with the national policy in this case was the lack of statute or legislation prescribing the language policy; this was a result of the Rwandan civil war which preceded the genocide. Furthermore, a national level peacekeeping or relief response was equally impossible, given that the governing forces were involved with the disaster, as an aggressor (Rawson, 2018).

3.5.1.1.4 Summary

In this case of the Rwandan genocide, there was a clearly justifiable language to use in any intervention or interaction with the situation, that of Kinyarwanda. This is evidenced by the speaking population being at least ten times larger than the second largest language, and it being a native pre-colonial language. Furthermore, there is potential issue of intra-language variation with Kinyarwanda, as there are seven recorded language varieties found within; however, little detail was presented in Ethnologue was provided other than that the varieties are close to each other; and additional searches failed to identify any evidence on the distances between the seven varieties.

3.5.1.2 Case Study Two

3.5.1.2.1 Background Information

The Second Congo War was a multi-nation conflict that began in 1998 and concluded with a peace agreement in 2002 (Cooper, 2013). The conflict started after 2 million Hutu refugees entered the Democratic Republic of Congo (DRC), those who fled the Rwandan genocide (Lopez & Wodon, 2005; Magnarella, 2005). In total the conflict involved seven nations and over twenty separate militia groups (Cooper,

2013), they were DRC, Rwanda, Angola, Burundi, Namibia, Uganda and Zimbabwe. The conflict resulted in at least 100,000 fatalities, although the precise figure is disputed due to a combination of limited, and biased reporting from the factions involved (Prunier, 2011). Following the peace agreements, the DRC then experienced the Pygmy genocide (2002-3), highlighting how volatile the situation remained after the war's conclusion (ibid). Despite the conflict involving multiple nations, the majority of the fighting occurred within the DRC territory (Reyntjens, 2009)

Prior to the commencement of the war, the linguistic landscape across the seven nations was diverse. Across the seven nations there were over three hundred languages recorded, with over half residing in the DRC itself (see table 9). There were several official languages across the nations. Three of the nations were monolingual in official policy, with Portuguese was the official language of Angola, French in the DRC, English in Namibia. Two nations were officially bilingual: Burundi, with Rundi and French; Uganda with Swahili and English. Two nations did not have legislated official languages during this time period: Rwanda and Zimbabwe. What was identifiable prior to data analysis was that there is no single language that held official status with all of the participating nations. The only evidence of a dominant language, that of English, is evidenced in the Pretoria Accord, which was the peace-treaty which concluded the war, as the text was presented and signed in English only.

Table 9: Number of languages per nations in the Second Congo War

Nation	Number of languages
Angola	38
Burundi	3
DRC	181
Namibia	28
Rwanda	3
Uganda	41
Zimbabwe	20

3.5.1.2.2 Results

For this case study, there were multiple target populations required for analysis.

There were seven nations in the conflict, which spanned five years, thus 35 values

were extracted from the World Bank repository. The population for each nation was

averaged across the five years, so that there were only seven find values for

calculating the total population. These seven values were then combined to form the

total population, on average, across the conflict overall. The Reach was then

calculated for each of the languages identified in the seven nations which held a

status of 1-3 in the EGIDS; which was 23 languages (see table 10). Where

languages were ranked with different status between nations, the higher status was taken forward, to promote the language status upwards in a minority-positive approach. The 23 languages accounted for just under sixty-six per cent of the total language use in the conflict affected population.

Table 10: Language Reach for the Second Congo War (all rank 1-3).

Languages with multiple statis entries, highest taken.

**Language Reach < 0.01%.*

Language	Status	Reach
Portuguese	1	8.66
Rundi	1	5.92
English	1	4.96
Swahili*	1	5.39
Kinyarwanda *	1	3.99
French *	1	2.42
Koongo	2	4.74
Lingala	2	1.21
Luba-Kasai	2	4.15
Bangala	3	2.07
Ganda	3	3.04
Kikongo	3	1.18

Kimbundu	3	2.37
Kituba	3	2.96
Shona	3	7.41
Umbundu	3	3.55
Afrikaans	3	0.05
Cokwe	3	0.27
Ndebele	3	0.92
Oshiwambo	3	0.27
Pidgin Bantu	3	0.18
Sango	3	0.00**
Lozi*	3	0.06

In this case study, the reduction step was taken. Whereby, only the top five languages, in terms of Reach, were retained for further consideration. All other language Reach scores were combined into a single category: Not Covered. This reduced the data set further to only including six values (see table 11). The top five languages accounted for thirty-two per cent of the total language Reach, a reduction of nearly thirty-four per cent from the twenty-three languages in the previous step. However, all of the languages included from this step were status 1.

Table 11: Top 5 language for the Second Congo War, ordered by highest Reach.

Language	Reach
Portuguese	8.66
Shona	7.41
Rundi	5.92
Swahili	5.39
English	4.96
Not Covered	67.66

3.5.1.2.3 Analysis

3.5.1.2.3.1 UNLP

A single language from the UNLP is part of the top 5 for this case study, that of English, which is a working language of the UN. However, this highlights how if a UNLP approach was taken then less than one in every twenty people would be able to access the communications. Whereas a policy with the largest five languages would still only provide language support of a third of the total population. If a single language approach was decided, the logical choice would have been Portuguese, where an average of two in every twenty-five people will be supported.

3.5.1.2.3.2 National Policy

All five of the languages that comprise the top-five group are national languages, with official status. However, the issue is that there is no single language that was found in all nations; English was the closest with five of the seven nations, showing that if a monolingual approach was taken whereby the most dispersed language was selected there would be substantive sections of the population being unsupported.

3.5.1.2.4 Summary

The diversity of this case study highlighted the difficulty that language policy makers face with international planning; even when only the majority and official languages are included there is still a significant issue of languages and language communities being unsupported. Another reinforced issue is how capping the number of languages to include in a policy can detriment the overall effectiveness of a policy, particularly when the policy is aimed at providing support to the total population in the area. In this case study, if the higher-status languages were all to be included, then two-thirds of the population would be supported, however, the logistic question of whether this approach is cost-effective in a disaster is unaddressed and will remain unaddressed in this thesis. A final observation from this data set is that the issue of language variation and Abstand is present in the dataset; across the seven nations there were three hundred and sixteen languages noted, however, there were more than one thousand dialects identified; and the extent to which these dialects deviate is inconclusive.

3.5.1.3 Case Study Three

3.5.1.3.1 Background Information

The Syrian Civil War began in March 2011, during the Arab Spring, with civil unrest and protests being held in the urbanized areas. The protests were supporting a democratic regime change for the national government (Zuber and Moussa, 2018) specifically the removal of the long-standing leader of the nation, Bashar al-Assad (Martinez and Eng, 2018) Multiple causes have been cited for the unrest, ranging from the widespread reduction of human rights (Martinez and Eng, 2018), to the political aspirations of the militia (Walther and Pedersen, 2020), to water drought (Karnieli, et al., 2019). Ultimately, no single factor can account for the entirety of events. The Syrian governments response to crackdown of the protests through violence springboarded further unrest (Zuber & Moussa, 2018) By July 2011, the Free Syrian Army was formed, with the purpose of fighting the existing Syrian Government for power of the nation. The conflict is still going (at the time of analysis). Despite attempts to broker a UN peace plan since 2012 (Tan & Perudin, 2019) the conflict escalated to the current status of civil war. The ramifications of this war include at least three million refugees being recorded by the United Nations Human Rights Council (UNHRC, 2022). The nation of Syria holds a single official language: Arabic, which is common across the nations in the Arab League.

3.5.1.3.2 Results

For this case study the population data for the nation of Syrian was collected from the World Bank repository; the average population through the conflict was calculated,

from 2011 to 2020. In total there were seventeen languages identified in the nation, however, only two languages which were awarded a status 1-3 using EGIDS (see table 12). Thus, the reduction step was not conducted, (similar to case study one). The top language accounted for around sixty-eight per cent of the total language communities; inversely, the second status 1-3, that of Standard Arabic, accounted for zero per cent. This situation is a result of standard Arabic being a taught language, and the thus it is not natively acquired, in contrast to North Levantine Arabic.

Table 12: Language Reach for the Syrian Civil War (all rank 1-3).

Language	Reach
Arabic, North Levantine Spoken	68.18
Arabic, Standard	0 (N/A)

3.5.1.3.3 Analysis

3.5.1.3.3.1 UNLP

Of the languages included in the Reach analysis, both were part of the UNLP, thus in this case study a UNLP approach would be suitable. Furthermore, when the languages with status 3 or higher were considered, the issue of language diversity and distance reappeared. There were six different varieties of Arabic accounted for, five which were colloquial Arabics, and the standard Arabic. If an Ausbaucentric approach to the UNLP policy was taken, then over eighty per cent of the population

would be supported (see table 13), however with Abstand this reduces to sixty-eight per cent.

Table 13: Language Reach of all Arabic varieties located in Syria during the Syrian civil war

Language	Reach
Arabic, North Levantine Spoken	68.18
Arabic, Standard	0 (N/A)
Arabic, Levantine Bedawi Spoken	0.54
Arabic, Mesopotamian Spoken	13.95
Arabic, Najdi Spoken	3.87
Arabic, North Mesopotamian	2.32
Other	11.14

3.5.1.3.3.1 National Policy

Similar to the UNLP approach, a national language policy approach also faces the same difficulties, that of Ausbaucentrism, and language distance. The de facto language of the state is North Levantine Spoken, this is evident due to the language's dominance across the total population, as nearly seven in every ten individuals in the state understand North Levantine Spoken.

3.5.1.3.4 Summary

This case study highlighted the issue of Ausbau-defined intra-language variation, and the extent to which Ausbaucentrism exists in an Arabic state. All of the possible approaches assessed would have been acceptable to use, as they would provide support the majority of the population.

3.5.1.4 Case Study Four

3.5.1.4.1 Background Information

This case study will be based on a theoretical future disaster, rather than a current or past disaster. At the time of this analysis, there were no major natural disasters reported or ongoing major conflicts that constituted a human disaster. This case study is to examine the policy approaches for pre-exemptive disaster situations. It is important to note here that this case study is not conducted as a prediction for any future disaster to occur in Ghana. This case study is included on the assumption that future disaster would satisfy all of the criteria. In that the disaster would be human led only occurring in the nation of Ghana, which would result in international intervention to provide a peacekeeping. Currently Ghana is a multilingual country with 81 language entries in Ethnologue and thus already satisfies two of the framework criteria: multilingualism & data availability. The official language is English.

3.5.1.4.2 Results

For this case study, the population of Ghana in 2020 was extracted from the World Bank repository, as the target population for the analysis. The Reach of Ghana's eighty-one languages were calculated. There were only seven languages which met

the criteria to be included in the Reach calculation (status 1-3 in EGIDS) (see table 14). The combined Reach of the status 1-3 languages was around seventy-three percent, across seven languages.

Table 14: Language Reach for Ghana languages status (all rank 1-3).

Language	Reach
English	24.97
Farefare	1.63
Mampruli	0.81
Akan	23.19
Ghanaian Pidgin English	12.74
Éwé	9.73
Hause	0.72

Following this, the reduction step was performed, reducing the languages to the top five (see table 15). This removed two languages and reduced the total Reach from seventy-three to seventy-two per cent. A minor reduction, and thus the top five selection here would still retain a majority of the population is covered.

Table 15: Top 5 languages for use in a language policy, ordered by highest Reach.

Language	Reach
English	24.97
Akan	23.19
Ghanaian Pidgin English	12.74
Éwé	9.73
Farefare	1.63
Not Covered	27.74

3.5.1.4.3 Analysis

3.5.1.4.3.1 UNLP

In this case study, speakers of the UNLP's working language accounts for nearly one in four of the population. However, if the UNLP approach was monolingual, then three in all four people would not be supported language wise. The diversity of languages, both in the top five and in the rest of the nation (all seventy-six) is still an issue for policy makers.

3.5.1.4.3.2 National Policy

This case study highlights how a single-language policy for nation can be supporting a minority, rather than a majority, especially when there is diversity across the language landscape. There is also a further issue of the dialects within the

language's not covered, so the difficulty in improving the national language official policy.

3.5.1.4.4 Summary

This case study highlights the disparity in linguistic landscapes between majority minority languages, given that seven languages in the nation account for three quarters of the nation's population. Therefore, the remaining quarter of Reach was from a combination of seventy-six languages, which averages as an estimated zero point three four per cent per status 4 and above language. Furthermore, this case study highlights how disparity can also appear with language size, and three of the languages account for over sixty per cent, due to them all being double-digit values.

3.6 Discussion

The results of this study address the three hypotheses' presented as well as informing an overarching additional issue, that of Ausbaucentrism (Tamburelli, 2014). The following discussion will be structured to firstly consider each hypothesis in order of appearance. Following which the issue of language demarcation and Kloss's (1967) Ausbausprache will be considered, particularly in the prevalence of Ausbaucentrism in African and East-Asian languages. In addition, there will also be a brief discussion of the thematic issues in current language policy: Colonialism; specifically, the acceptability of selecting the languages enforced by colonial powers historically for use in post-colonial states and Multilingualism; notably accounting for a diversity of languages in the field, (see *2.3.2.4 Criticisms of current international language policies*). To remind the reader, the three hypotheses' were:

1) Selecting languages for use in an emergency, via a Reach approach, will enable greater language support from the offset of policy deployment than arbitrarily using the languages of the UNLP; as a Reach approach will provide an element of bespoke-ness to the situation and will account for the regional languages at use in the disaster-impacted population.

2) A selection of the largest 5 languages in terms of Reach will not be sufficient in providing language support to the majority of the population, due to the prevalence of language diversity and minorities globally.

3) Assessing Reach will provide alternative choices for language policy formation than using current international and national policies as the primary guidelines, as an access approach will account for the deploy-ability of a language.

3.6.1 Hypothesis 1

The findings of this study support the use of a Reach approach in emergency language policy as opposed to using the UNLP as the de facto guideline set of languages. The Reach approach identified different lists of which language to use for each of the four case studies analysed; with all containing at least one language which is part of the UNLP. This result may be taken to support the use of the UNLP as the de facto language policy guidance, given that the UNLP was, in part, applicable in the cases. However, only one of the four case studies featured two languages from the UNLP, with the others containing just a single language (English). In fact, the supporting evidence for using the UNLP is predominantly due to the presence of English in Africa and Arabic in the Middle East.

In the case of the Syrian civil war, the use of the UNLP would have been appropriate, due to Arabic featuring in the UNLP and the majority of the Syrian population being classified as Arabic speakers. However, the colloquial Arabic variation within the nation is another element which may cause communication difficulties; given that language variation resulting in low intelligibility within the idealised Arabic language is an established element (see *2.3.4.5 Intelligibility in Arabic*). This issue is due to Ausbaucentrism, which is an overarching issue and will be addressed as a separate section following the discussion of the hypotheses.

Furthermore, the results indicate support for the adaptability of the Reach approach, as the lists themselves did not overlap, in that each list by-and-large did not contain the same languages, with the only exception being English. It is worth noting that French was reported in multiple case studies but did not get included in the Reach final list of languages to use due to a lack of speakers in the overall situation's, therefore, the language could be used, but it is not advantageous when compared to the other languages in the case. Additionally, the results of this study identified that the application of the additional three UNLP languages (Spanish, Russian and Chinese), in the case studies would have been redundant as there were no speakers or speech communities identified, therefore, supporting that arbitrarily applying the UNLP is not appropriate in the field.

3.6.2 Hypothesis 2

The approach of using the top 5 languages for an emergency language policy is particularly supported by the results of this study. There were three case studies where the top 5 Reach languages would have provided language access to half of

the total population in the disaster event. The single case where language access did not breach the threshold of a simple majority (as in greater than 50%) was the Second Congo War, with an estimated third of the population receiving language access using the policy alone. The defining distinction between these two groups is the size of the conflict itself, as the case studies which supported a top 5 language list approach (which was achieved with a total Reach score of over 50%) were all single nation events, with a single society and government involved. Whereas the Second Congo War was a multinational conflict, with multiple governments and societies involved overall (Cooper, 2013). The size differences were also found in the population size of the conflict and the extent of linguistic diversity recorded in the conflict. There were multiple national level languages across the disaster zone, as each nation's language policy varied from their neighbours, there were languages which held high EGIDS status in one nation but held lower status in another nation: highlighting the variation of language access across a multi-national situation.

Overall, the results support the rejection of arbitrary language policy for multinational conflicts, as the application of the UNLP in the case studies would have languages deployed in situations where there were no speakers, making the deployment itself redundant and illogical. Additionally, the application of a standardised list of languages would not account for the linguistic diversity across multi-national situations, as there is no guarantee that languages which were spoken predominantly in a single state will be appropriate for use (as in accessible) in neighbouring states, even if the language is present in both states. This is evidenced through the conflicting language status's reported in Ethnologue, as the status of a language is

based on the use of a language in a nation exclusively; for instance, French is a global lingua franca (score of 0 in Ethnologue), however, in Africa, there are nations where the status is lower (as in 3 in the DRC), due to a lack the infrastructure and usage of the language beyond the officialdom for the state. Furthermore, these results do not support the absolutist altruistic approach to emergency response, whereby all individuals must be fully catered for equally in the language support provided, so that all languages are treated equally and there is no unfairness or discrimination occurring from the policy itself (O'Brien et al., 2018), this rejection is motived by the feasibility of the approach, given that language diversity is a major issue facing policy makers (TWB, 2015). Additionally, the results reject the notion that all emergency and disaster plans should consider equity as a minimum standard from a human rights perspective (Fierros et al, 2017), again on the grounds of feasibility, or lack thereof, at present. For clarity, the Second Congo War case study contained information on over 300 languages existing within the conflict zone, a volume of languages which currently is unrealistic to manage in an all must be accounted for equally from the outset approach. Furthermore, of the 300+ languages identified, the vast majority (as in 300+) were minority languages with remote and spare speaker populations, as such, the effort required to provide support brings with it the question of whether the benefits are objectively worth the concerted effort, especially when it comes to the languages which are not either standardised or written, as these are two primary features required for translation services to operate effectively (Larkin, et al. 2007 & O'Brien et al, 2018) It is worth acknowledging that this study supports Jones & Askew (2016)'s notion that equal language access

should be the aspirational goal, whereas the operational goal should be to cater for the majority impacted instead.

3.6.3 Hypothesis 3

The results of this study support the calculating of language access as the primary criteria for forming emergency language policy. The core evidence for this position is found in the Ghana case study (to recap, this case was predictive, as was designed to explore the languages to use for an preparing language policy and resources in a nation where there currently are no active human disasters with Reach calculated using up-to-date data), as using both the UNLP and the national language policy would advocate for the singular use of English as the language to use; an approach which would provide some language support, as the language access score for English was just-under a quarter of the population. However, when Reach is applied to the situation, the results support retaining the national language (English) alongside two additionally languages in a trilingual approach instead which would cater for an estimated 60% of the population. This highlights how using Reach does not necessarily result in the discarding of (inter)national languages, rather the approach can support the expansion of language policies whilst retaining the current infrastructure.

The primary advantage of using a Reach approach is that the policy calculated indicated maximum potential for language coverage, prior to the implementation itself; as a Reach score are the maximum coverage of a region that can be achieved when using the specific language; therefore, policies using Reach themselves have a ceiling of effectivity inbuilt. It is improbable, but not impossible, for the effectively of a

Reach policy to be greater than the score of Reach itself, as to achieve this a language policy would have to provide communication support for individuals who do not speak the languages included in the policy. In the event that a language policy provides a greater percentage of the population catered for than the maximum Reach score calculated would likely be a result of either high intelligibility between the languages in the area, with speakers being able to understand parts of the messaging delivery in the languages selected for use; as the continuum of intelligibility is not binary, and allows for partial communication between languages at varying degree of linguistic distance (Gooskens & van Heaven 2021), however, this is an element which cannot currently be considered with Reach calculating without additional information regarding the status of intelligibility between languages in the case studies studied.

3.6.4 Ausbaucentrism in African and East-Asian Languages

Throughout the study, the data collected from Ethnologue contained information discussing the language varieties found within the languages recorded in Ethnologue's database. This study's design was to collate the available data on established language entries from the linguistic database, so during the data processing and calculating, the linguistic distance of intra-language variation was not considered. This limitation was designed to restrict the scope of the data and to provide an objective and measurable response to the research question. Although this did achieve the task set forth i.e., sufficient data was retrievable and the Reach scores were calculatable, the issue of Ausbaucentrism cannot be ignored in discussing the situation of language policy formation, nor can the issue be side-lined

when considering communication-based language policy. The issue of Ausbaucentrism is prevalent in all four case studies conducted in this study as each used Ethnologue data entries that contained Ausbau languages with multiple languages varieties included in the language entry with little information relating to the any Abstand measurements.

As a recap, Ausbaucentrism, is when the defining criteria for language demarcation is based exclusively with Ausbausprache features or characteristics (Tamburelli, 2014). This includes classifying two or more linguistic varieties as being part of the same Ausbau language, which is typically done by political authorities to support the idea of unity between the speech communities (Wright, 2016). The application of Ausbau, in itself, is not an issue as the socio-political status of a language is an established vital part of language (Kloss, 1967). The greater issue is found when the ability to understand or communicate between speakers of the same language is impaired or restricted because of the Ausbau-led definition of the linguistic varieties as a single language; an issue which has been reported in multiple contexts and languages (with Chinese (Tang & Van Heuven, 2007, 2009 & 2015), Scandinavian languages (Kürschner et al., 2008; Schüppert et al., 2016), Italian languages (Tosco, 2008; Tamburelli, 2014; Tamburelli & Tosco 2021), Madagascan languages (Bouwer, 2007) and Arabic (Zbib et al, 2012; Kwaik et al. 2018a)). It is assumed that speakers of the same language will be able to communicate with one another, however Ausbaucentric approaches do not account for the ability to communicate within the community as a factor for classifying a language (see Leonardi & Tamburelli, 2021 on a similar issue). This results in situations where linguistic diversity and distances

are ignored officially but these differences are a significant issue for on-the-ground discourses, as reported in TWB (2017a), where communication was hindered or impaired, despite the assurances from political sources that the speakers all speak the same language. These situations highlight how the political perspective on the situation, i.e., that there are no issues, is starkly different to the reality (Jones & Askew, 2016).

The results of the Syrian Civil War case study expose the Ausbaucentric approach towards defining Arabic linguistic varieties, particularly with the differences between colloquial and standard forms. If an Ausbaucentric approach is taken, the language Reach would exceed 90% of the population in the disaster zone, which explains why previous policies took this to be a good solution. However, when considering the gap between an Ausbaucentric approach and actual communicative effectiveness (using Abstand approach, following the same work of Kloss, 1967), a very different picture emerges. For example, Ethnologue demarks multiple distinct Arabic languages in replacement of a single Ausbau language, calling into question the supposed 90% Reach. The justification for demarking multiple Arabic's is rooted in Abstand, as the distance between the colloquial Arabics is considered sufficient to treat the language varieties as distinct in Ethnologue's database, which is the minimum requirement for presenting a language as distinct in the database (Eberhard et al., 2022). However, it is important to remember that re-classifying Arabic is not a nuanced approached, as there have been multiple demarcation of Arabic presented in the past. Notable classifications include the generalist approach (Versteegh, 2014), which splits Arabic into two distinct languages based on the geo-cultural (Al-Wash 2016) and anecdotal

linguistics evidence (Trudgill 2009; Bassouniey, 2012; Horesh & Cotter, 2015), which are the Mashriqi (Eastern Arabic) and the Maghrebi (Western Arabic). A classification which clashed with Ethnologue's depiction of the status of Arabic, that of there being multiple Arabics as opposed to a binary set, as there is disagreement between both classifications as to which language varieties are in fact distinct languages.

Another classification is Zaidan & Callison-Burch (2011, 2014) who supported re-defining Arabic into five distinct languages, which was supported by Zbid et al.'s (2012) study on the calculated lexical and morphosyntactic differences between the usage of the languages on social media; within this reclassified system, the Syrian Arabic varieties were classified as being Levantine Arabic. The conflict is direct with Ethnologue, which was classified both by Zaidan & Callison (2011) and Zbid et al. (2012) as defined Levantine as a single language, yet Ethnologue separated it into two different languages (North Levantine which was present in Syria and South Levantine, which was not present. Furthermore, as well as distinguishing several other colloquial Arabics as separate languages (such as Mesopotamian Spoken). This conflict between sources on the demarcation of Arabic exposes a research gap within language mapping, as there is active disagreement between Abstand-method linguists as to which Arabic colloquial varieties are in fact distinct languages. Conflicting reports of language status is not a new phenomenon in linguistics (see *2.3.1.4 Abstand and Ausbau*), however, it is evident that Arabic is considered one language only from a Ausbaucentric perspective, which does not help when gauging its potential for communicative purposes, since Ausbaucentrism does not consider

intelligibility as a criterion. Efforts to provide Abstand demarcation are still developing with the complex situation that is Arabic colloquial variation.

In this study, there were two major issues rooted in Ausbaucentrism: data inconsistency & criteria inconsistency. Overall, the pejorative issue is the lack of application of Abstand measures, such as intelligibility, within linguistics and Ethnologue in defining a language. As, throughout the case studies, the Reach language lists contained languages which hold the characteristics of Ausbau languages such as English (Trudgill, 2002; Sussex, 2004), Kinyarwanda (Nkejabahizi, 2007; Nassenstein, 2019), Akan (Kiyaga-Mulindwa, 1980; Guerini, 2008) and Arabic (Zbid et al, 2012; Manfredi & Tosco, 2018) as there are multiple varieties within each language , which are disputed to being intelligible between each other, thus classifying as Ausbaucentric due issues of low intelligibility. The use of Ausbau-defined languages itself is an issue, as the approach needed for disaster communication requires communication ability and intelligibility to be considered as primary criteria (TWB, 2017a); the use of an Ausbau language does not guarantee that all the speakers identified within the group will be able to passively receive information using the language, let alone participate in discourse to add to the conversations.

The issue of data inconsistency is the level of information provided for each language entry, specifically whether there are linguistic varieties within the language's described in each entry. For instance, the entry for Chokwe in Angola describes the language as having over ten different names depending on the location of the population, however the entry does not contain any information as to whether the

alternative names are a result of language variation. Fleisch (2009) describes

Chokwe as containing multiple linguistic varieties which vary greatly from standard

Chokwe, with these variations being considered as still part of the language of

Chokwe, as common Ausbaucentric approach. Another example is Diriku in Namibia,

which is recorded to have multiple alternative names (such as Gciriku, Manyo and

Shimboyedu), yet there are zero known language varieties identified within the

overall language. There is an issue here, as Hammarström (2019) defined Manyo as

a mutually unintelligible variety of Diriku, which disputes Ethnologue's entry in two

ways; firstly, that Diriku is also Manyo and secondly that there are no alternative

linguistic varieties labelled within Ausbau-defined Diriku language. It is important to

also acknowledge that all the examples used in this section thus far have been part

of the Bantu language family. Bantu languages are found predominantly in central

and southern Africa and the issues with data are indicative of the imbalance of

language investigation by geographical region, as African languages have been

regarded as under-researched in comparison European or Asian language families,

such as Proto-Indo-European or Semitic (Williams, 2011; Johnson, 2013). As this

study exclusively explored situations involving African, or African adjacent nations,

this discussion cannot infer that the issue of data is an overarching criticism of all

language data collection and management; rather this commentary is limited to the

languages of the African continent and the Arabian Peninsula.

This lack of available data exposes an issue for not only Ethnologue, but also for

disaster language planning, as the lack of intelligibility data for language demarcation

and Abstand led comparisons hinders efforts to account for the appropriateness of

using a language in an emergency; specifically, the lack of data supporting the application of Abstand approaches is of paramount importance. Each Ethnologue entry is written with the information that Eberhard et al., (2022) has identified from across the literature, and across the literature there is an imbalance of information regarding all languages equally. Additionally, there is a secondary level of data to consider too, that of locationality of the language in question, as the Ethnologue entries do not report on language in a global perspective, rather the entry provides the localised information on a per country basis, as such the information requires both the information the language itself, such as speaker numbers, and on the socio-linguistic factors, such as language vitality status and the acceptable contexts of which the language is used in. The two-fold localisation is a primary issue across language planning data, as the collection of data is difficult when the linguistic and physical landscapes are diverse. For instance, in African countries a high level of linguistic diversity is established (Djennadaal, 2008; Williams, 2011), and the speaking communities are both in differing states of human development as well as geographic, with some communities being isolated in remote regions of desert or jungle (Nettle, 1996). The issue of data inconsistency is logical, given the depth of diversity and the logistical difficulties in reaching isolated communities, however, the extent of the problem is evidenced significantly in Ethnologue's data entries for African nations; to a degree greater than initially expected upon the commencement of this study.

Furthermore, there is the other primary issue, that of criteria inconsistencies with the demarcating of both languages and the varieties within the languages themselves.

Unlike data inconsistency issues, which are logistic as well as linguistic, the decision of appropriate criteria, i.e., which features to investigate or measures to use, is an issue exclusively within the remit of linguistics and sociolinguists. Across the data available from Ethnologue used in this study, there were multiple criteria for language demarcation presented, with a lack of qualifying threshold for when the linguistic distance constituted a language variety being classified as a distinct language. The primary tool for comparing inter-language variation in Ethnologue was lexical similarity, between the standard or dominant language variety from an alternative language variety; a measure which was marked as a percentage of shared lexicon. However, there was another method deployed in tandem in some cases, that of intelligibility and mutual intelligibility. The issue here is the inconsistency in which both measures are deployed. For instance, there are languages (such as Bwa, Nyanga and Lombo) being considered intelligible with a percentage of understanding presented as evidence, however there are others presented as intelligible to all related linguistic varieties without any supporting evidence (such as Mongo-Nkundu, Bamwu). This supports the notion that language mapping data is flawed, as methods deployed to mark language distance and variation are inconsistent (a view shared by Spolsky (2011), Bouwer (2007), Tosco (2008 & 2011); Kwaik et al (2018b) and Leonardi & Tamburelli (2021)); in fact, this thesis supports that the deeply rooted issue is still ever-present in modern language mapping.

Additionally, there are underlying gaps which support that some language entries should be split into two or more languages from the single Ausbau language (such as Kituba & Koongo). Another issue is the threshold in lexical similarity required to be

considered a single language, as there are entries that record a similarity score below 75% as being particularly unintelligible, and therefore possibly different languages (such as Budu) yet there are others which score below 75% and are considered as part of the same language (such as Ganda in Uganda with 55% similarity with Kooki, a language variety recorded as being a sub-language of Ganda).

Mutual intelligibility is also used as a method for referring to inter-language distance and connection, however, what is not addressed is whether the languages marked as mutual intelligible are considered so from Ausbau-rooted or Abstand-rooted evidence. For instance, in Uganda Kenye is recorded as being mutual intelligible to all known sub-varieties of the language (Siginyi, Gwere, Lamogi and Lukenye-Lukooli); the issue lies in Ethnologue recording many of the sub-varieties as individual languages' (Lamogi, Siginyi and Gwere).

This leaves the question of whether the speakers within these communities will be able to understand one another whilst also being able to accurately communicate between themselves, unanswered. Deploying the languages selected using Reach could result in the same communication issues and difficulties in the field as the current approaches as well as the UNLP; this is a major concern and issue that needs exploring in greater detail and with further research, as the objective measure of Reach may be partially fallible due to the reliance on subjective or inconsistent data entries in Ethnologue. However, it is also important to recognise that Ethnologue is the primary database for language mapping at this time, and that alternative databases' preliminarily explored for use in this study were rejected due to a lack of

available data for minority languages. An example is the World Atlas of Language

Structures (WALS) (Dryer & Haspelmath, 2013) a database which focuses on

mapping established language features and characteristics, such as phonological

inventories and morphological systems, thus the required data, i.e., speaker numbers

and localised language vitality status, was not available from the WALS as a source.

Therefore, the criticism of Ethnologue, is aimed towards improving the system, rather

than advocating for the avoidance of the database itself; as the project's size and

scope provided sufficient information for the development and application of the

Reach and Access approach proposed in this study.

3.6.5 Colonialisation

The findings of this study also inform the debate concerning post-colonialism and

language policy on the international diplomatic stage. Wright (2016), noted that

colonial languages retain global dominance in terms of use in diplomatic and

international discourses, such as treaties and legally binding agreement. The debate

currently focuses on the appropriateness of using colonial languages in postcolonial

settlement and nations (Hyunjung & Ryuko, 2008; Fitzpatrick, 2018; Mir, 2019) and

from this study's results are a partial response can be formulated. The results of this

study indicate that the use of international colonial languages is not without logic,

particularly that in when all case studies were analysed, and a top five reach list

compiled; three quarters of all lists contained a colonial language. Therefore, if the

appropriateness is measured by whether the language can be used for language

policy, and in doing so provides access to greater portion of the population (in

perspective to other languages in the region), then this study would argue against

efforts to remove colonial languages arbitrarily. In three of the case studies, if the colonial languages were rejected, for being colonialist, then the Reach scores would drop, as the policy would deliberately exclude all individuals who contained to use the colonial language as their primary language in the post-colonial era. The application of colonialise language removal would hinder the policy in achieving the overall aim, that of providing maximum language support in an objective manner (Gerver, 2021a), not least because the removal of colonial languages would be politically motivated and arbitrary to the overall situation in direct contrast to the objectivity clause, but also because the rejection of colonial languages would exclude the sizeable groups in the situation, which ultimately undermines the altruistic approach of helping as many as possible, an approach which governs all disaster management (Maestri & Monforte, 2020).

Furthermore, the results of the study do not support efforts to remove colonial languages entirely for an additional reason, as the removal of any higher status (in terms of EGIDS) language from an emergency language policy should not be done unless there is an alternative equal Reach status (here referring to the EGIDS language status threshold (national or above) required to be considered in a Reach list) language that can replace the colonial language. As language policy for use in emergencies requires existing resources, and language as a resource, to be considered. This study was focused on investigating whether language could be deployed in an emergency, in its current state; therefore, the removal of any languages which already enjoy social political prestige, teaching resources and/or high density of speaker numbers in the population, cannot be advocated. To remove

a language from a prospective policy based on political factors, such as postcolonial pressure, would undermine the overall effort of quantifying language policy decisions in a way to remove any socio-political interference; and as such is not supported both by the evidence of the study and by the overarching position taken in this thesis.

3.6.6 Diversity and Multilingualism

Additionally, the case studies selected provided a scope of disaster types, in terms of land use and human diversity. The results of calculating reach highlight the difficulty faced in all language policies when diversity in terms of language use is considered; particularly the overarching issue of planning language resources policy to cater for minority language speakers. An issue that is not new from this study, however the results of the Second Congo War exposed the usage of a single top-down policy approach for language use in an emergency has, in itself, pitfalls when the community is diverse, a notion supported for general (or peacetime) international language policies (Johnson, 2013; Wright, 2016; Kazeem & Suleiman, 2020). In the case of the Second Congo War using a Reach language policy would still not cater most of the population in the conflict; however, it is important to note that using the United Nations language policy as the framework would produce worse results, with less than 10% of the population coverage. Overall, the study supports efforts to quantify language policy for interventions, however the results support TWB's (2017a) notion that linguistic diversity is still an overarching issue which warrants further research and investigation; with the aim to provide language support and access to the minority populations.

3.6.7 Responding to research question 1

(1) Are international language policies suitable for use in emergency disaster language policies?

When considering disaster language policies, the requirements are that the languages selected are deployable, acceptable and preparable. The requirements of each are detailed in this chapter, as well as the application of this criteria to historic and current cases of disaster in using a new mechanism (Reach). The findings support the viewing of the UNLP, the world's largest current international language policy, as potentially acceptable for use in future disasters, as well as additionally particularly successful in historic and current disasters. The tentativeness of considering the UNLP as usable is rooted in the three criteria for language selection. With the first criteria, deployability, the language's currently in the UNLP are more favourable for use than national languages, and especially more favourable than minority languages. There are established resources for translation, teaching, and usage of the six languages within the UNLP, and as such, there is a global advantage for the languages, as additional language would require extensive effort to develop the resource stockpiles to match the current six UN languages. The limitations for using the UNLP is the intransigence within the global politisphere towards adapting the policy to be more representative of the current status quo of languages in the world (d'Almeida & Octu-Grillman, 2013). Additionally, there is the altruistic clause within the UN charter itself, whereby all languages are acceptable within the UN's discourses and discussions; this is designed to support the fundamental language rights however, it also proposed that situations where a

language cannot be used as not only problematic, but potentially in breach of global human right's law. In a disaster situation, the focus is on maximising the reduction of harm faced by the greatest volume of people, therefore the UNLP's focus on human rights can be seen as particularly counter-intuitive to the utilitarianism required in large-scale disaster policies.

The case study analysis highlighted that the UNLP is not a catch-all policy due to Ausbaucentrism and multilingualism. In most cases the use of the UNLP alone would not have provide sufficient support for the populations impacted, however, adapting the policy to include additional languages found within the target region was found to improve the Reach of a language policy substantially. In the case of Rwanda, the inclusion of Kinawarwya would have encompassed the nigh-all the population in the Genocide; however, the use of only the UNLP would provide language support to less than 10% of the population. Furthermore, in the Ghana case study, the inclusion of Akans would have doubled the Reach for a language policy to an estimated 50%.

In the Syrian case study, the UNLP was most successful, the issue of Ausbaucentrism was identified. Using Arabic would provide support for most of the population, however, the data entries from Ethnologue identified multiple distinct Arabics, where are classified as languages. This exposed the reliance within the UNLP on Ausbau-defined languages, and the disregarding of language variation and communication barriers within the same languages. The rational for the existence of the Ausbaucentric approach retained in the UNLP is rooted in the historic geopolitics which were used to designate which languages to use. As such, there is opposition to the amendment of the UNLP in the formal use within the UN operations and

activities. However, when reassessing the potential Ausbaucentrism within the languages of the other three cases studies, further issues and discrepancies were identified related to which languages were demarked using Abstand methods and which by Ausbau, predominantly, the lack of evidence either way. Therefore, the application of the UNLP for a disaster policy would require a further investigation as to whether the local realisation of the UN language would be intelligible for an outsider speaker.

With multilingualism, the case of the Second Congo War exposed the issue with using a limited size language policy, as well as the issue of minority languages found within a target region. The deployment of the UNLP in this case would have provided support for around 5% of the population, which is lower than the estimated 8% reported for the Southern European Crisis language response. Furthermore, there were over 200 languages recorded in the region which in total, amounted to less than 1% of the total population, and providing equal support for all languages is currently unfeasible. It is worth noting too, that the use of the top 5 languages, by Reach score, would have provided support for around a third of the population.

Overall, the UNLP can be used as a guideline policy for the required resources for deploying a language policy, and it is the position of this thesis that the UNLP can be applied for the initial stage of a disaster. Despite the discussed issues with the policy, the status of the six languages, in terms of deployability and resources, is a scaffold to build upon for future policies. Therefore, in summary, using current international policies is an acceptable start-point for language policy planning, but it is not

acceptable for use without amendment to suit the linguistic landscape of the target region.

3.7 Conclusion

The findings from this study indicate support for a reassessment of language policy in situations of emergency and disaster, as there were significant gaps in language policy coverage when the language policy is arbitrarily connected to international language policies. For instance, using the UNLP as the default language policy for three of the cases studies would have produced situations with language accessibility being obtainable for less than half of the total population. Furthermore, these results support the hypothesis that using the objective measure of access provides a detailed base-line mechanism for policy decisions in further language policy use; specifically, the issue of which language to use in a policy. Additionally, these results provide evidence of extent to which Ausbaucentrism is the de facto status for language demarcation, as well as, providing evidence for the severity of issues when using Ausbau-defined languages in active policy for emergency context. It is also worth highlighting the inconsistency of language demarcation techniques within language mapping, specifically in relation to the measuring of mutual intelligibility between two or more linguistic varieties.

Overall, the issue of Ausbaucentrism in both Ethnologue's database and in the African languages are significant obstacles which require further attention; however, this thesis will not progress further with the specifics of both Ethnologue's database and African languages (such as Bantu languages). The issue of Ausbaucentrism in

East-Asian will be progressed further, specifically the Ausbau language of Arabic.

This progression is warranted from the diversity recorded in the Syrian case study,

where multiple Arabic varieties were recorded as being distinctive languages,

coupled with the literature having recorded issues of linguistic variation, varying

levels of intelligibility between Arabic across a number of demarcation vectors (such

as Mashriqi vs Maghrebi).

Chapter 4: Exploring Mutual Intelligibility in Mashriqi Arabic's

This chapter contains a study which investigated the intelligibility relationship between Nijari Arabic (as a Gulf QA), Cairene Arabic (as an Egyptian QA) and MSA (in context) which replicates, in part, the demands of the field describe in TWB (2017a). The aims of this study were to explore the status quo of languages intelligibility within Arabic speakers, in a comparable situation to the field, so that the results can develop our understanding of the real-world issues reports and ascertain whether intelligibility can account for part of the lack in communication between refugee's and aid workers. Intelligibility was measured through a picture selection task, with the accuracy of response, and the speed of response used as the factors to score the continuum between the three Arabic's used. This chapter will also detail the development of stimuli and experimental design for the calculating of intelligibility between multiple language varieties with a continual contextual factor (which in this case was cognitive load to stimulate a form of stress). This study proposes an Abstand-led approach to quantifying the language distance between two Mashriqi QA and the standard international Arabic of MSA; to allow for the objective consideration as to whether a language variety is distinct enough to be classified as a language.

4.1 Research Question

This chapter will explore the second research question proposed from the literature review: are the Mashriqi colloquial & standard Arabic varieties found within the diglossic continuum of the Kingdom of Saudi Arabia sufficiently intelligible, that it does not matter which variety is used in stressed contexts, such as in emergencies?

4.2 Aims

This study aims:

- To investigate the extent to which the intelligibility between Arabic varieties inhibits communication between speakers in stressed discourse, if any.

- To explore if prioritizing a specific type of Arabic (QA/SA) would improve general intelligibility and accurate communication in stressed contexts.

- To provide empirically evidenced insight into the status quo of communication between Mashriqi Arabic varieties in stressful conditions.

4.3 Hypotheses

In this study there are three hypotheses, motivated from the state of the art.

1)   The local variety of QA will be the best result for both accuracy and response speed, in comparison the other Arabic varieties exposed, (as the local variety is the primary language for the listeners and therefore is the most established and exposed form of Arabic to the participants).

2)   Colloquial Arabic varieties will be responded to with greater accuracy and at a faster speed than standard Arabic, (as colloquial Arabic's share linguistic similarities features but these similarities are not shared between colloquial and standard Arabic's; therefore, the processing will require greater effort which will slow down responses.)

3)   Modern Standard Arabic will not function as the middle variety in terms of performance in responses, (in a rejection of the Ausbau centric idea that standard

Arabic is the go-between for communication between all Arabic speakers regardless of their origin or their native colloquial varieties spoken.)

4.4 Participants

Participants were recruited via the internet using the investigators personal networks, such as LinkedIn & Reddit. Participants were asked to select, from a drop-down menu, where they were from; there were multiple regions possible to select, but only those who selected Saudi Arabia were included in the sample. The design of multiple options was used to elicit honest answers, as there was no indication that only Saudi Arabians were required for the task. All participants were asked to provide information on their age & sex via selecting the appropriate bracket, again from a drop-down menu.

4.4.1 Pre-testing Questionnaires

To obtain the demographic information & consent, participants were tasked with completing a questionnaire before the commencement of the main task. These questions were designed using Gorilla Experiment Builder (2020) which allowed for interactive questionnaires with drop-down menus providing a set list of options. The questionnaire allowed informed consent to be obtained, with participants giving consent in a tick-box; completion of which was a requirement to continue. Below the tick-box was a list of demographic-based questions, each containing a dropdown menu for answers. The answers were fixed to allow for an automated system to sort the participants; if an answer was selected which was unneeded, then the participants would be re-directed to a message informing them of quota filled for

participants with the selected demographics. This allowed the building of the target groups to be conducted without interference or intervention from the researcher. At the end of the questionnaire screen, the participant was reminded, via another mandatory tick-box, to ensure that their sound was turned on.

The questions and responses options were: (rejected responses are flanked by [ ] for the readers clarity)

Sex – Male/Female/[Other]

Age – 18-24/25-32/33-40/[41-50]/[51-60]/[61-70]/[71+]

Region of Origin – [Egypt]/[Lebanon]/Saudi Arabia/[Tunisia]/[Other]

Forty-three participants (M = 22 – F = 21) in total took part in the experiment, these 43 identified across two age groups; 18-24 and 25-32 (see table 16). All participants were based in Saudi Arabia when completing the study.

Table 16: Age and sex grouping of the participants.

	18-24	25-32
Male	13	9
Female	10	11

4.5 Stimulus Design

For this experiment, two sets of stimuli were developed for a language task, and two set produced for two cognitive load tasks. The first language task set was a set of

regular simple sentences which could be then translated into three Arabic varieties (Nijari Arabic a Gulf Arabic; Cairene Arabic an Egyptian Arabic, and Modern Standard Arabic) and subsequently recorded. The second language task set was a set of images which were related to the sentences selected from the first task set. In total there were 39 sentences used experimentation phase of the study, recorded in the three Arabic varieties for the first language task set; and there were 156 images created for the second set. The cognitive load tasks required set of mathematical equations of varying difficulty for the first task, and block images of colours at varying saturation level for the second task. For the first cognitive load task set, there were 39 equations produced, and for the second set there were 156 images produced. The following section will outline the development of the stimulus whilst also providing examples of the final products.

4.5.1 Sentence Extraction

For this study, a set of sentence-length stimulus were developed. The base-form sentences were initially extracted from Kalikow et al.'s (1977) SPIN test, which were developed as a tool for assessing speech reception in individuals with hearing impairments. Kalikow et al.'s (1977) study produced a set of 400 sentences, ordered into 8 lists for use in assessments. The length of each sentence ranged from 5-8 words each, with a between 6-8 syllables in English.

There were two distinctive types of sentences produced for the SPIN, based on a distinction in semantic predictability of the final word within the sentence (also known referred to as the 'key word'). This predictability was related to whether a speaker could accurately guess (or predict) the key word if there was interference (or excess

noise) in the key word when heard. Sentences of High Predictability (HP) were designed that the listener would be aided by "the syntactic, semantic, and prosodic cues available in the sentence" (pp. 1339), thus resulting in high levels of accurate prediction of the key word. Whereas the Low Predictability (LP) sentences were designed for weaker assistance in predictions, as the listeners "must depend … on acoustic properties and lexical information concerning the key word itself" (Kalikow, et al. (1977) pp.1339).

The SPIN lists were also balanced in the final presentation for "intelligibility, key-word familiarity and predictability, phonetic content, and length" (pp. 1337). With familiarity, Kalikow et al. (1977) used the Thorndike-Lorge list of word frequency (see Thorndike & Lorge, 1952); deciding to only use words which had a frequency of 5 to 150 occurrence per million words. Whereas, for intelligibility, phonetic content, length & predictability, Kalikow et al. (1977) tested these measures themselves across a range of experiments. The overall study result was a collated set of sentences arranged into lists, which were all assessed by a breath of criteria. Below are two example HP sentences from SPIN (see Kalikow et al., 1977: 1348-9, for the full lists).

Form 2.1: 2. She made the bed with clean sheets

Form 2.1: 7. He caught the fish in his net

For this study, the 200 HP sentences were extracted for use, a volume greater than required, as such the list of extracted sentences were then reduced using a two-step filter. The first filter applied was regarding the clarity of the key word, when the key word is presented as a single static stock image; and the second filter being removal

of idioms. The next table (see table 17) will provide example LP sentences, as well as example HP excluded from the extracted list by the two-filter criterion.

Table 17: Example SPIN sentences excluded

Exclusion criteria	Example	Issue with example sentence
Low Predictability (LP)	'Jane has a problem with the coin'	The key (final) word, *coin*, is not semantically inferred by the preceding words in the sentence
Clarity of Image	'I can't guess so give me a hint.'	The key word is part of an abstract concept, which was deemed too difficult to visualise in a single image. So, whilst the key word may be predictable when spoken, a graphical representation of the key word is unfeasible.
Idiom	'He was scared out of his wits'	The intended meaning does not correlate with the literal meaning. Translating idioms between English and Arabic is problematic (Al-assaf & Abdulaziz, 2019; Banikalef & Naser, 2019).

After the filtering process was completed, there were 72 possible sentences from the original extraction list remaining. To produce a final sentence list, the number of sentences was reduced to a list of 40. This reduction was performed by randomly splitting the 72 sentences in 8 equal-sized groups (8 groups x 9 sentences). The first five sentences in each group were then selected for the final list (8 groups x 5

sentences = 40 final sentences). This list was the final collation of the extracted SPIN sentences into an experimental list of sentences.

4.5.2 Translation

The list of 40 sentences were translated from English into three Arabic varieties: Cairene colloquial Arabic (Egyptian), Nijari colloquial Arabic (Gulf) & Gulf-accented Modern Standard. The translations were conducted by two female postgraduate linguists who grew up in the national countries of at least one of the target Arabic varieties; one Egyptian from Cairo who spoke Cairene Arabic & one Saudi Arabian from Riyadh who spoke Nijari Arabic. The Egyptian student translated the sentences into Egyptian, whereas the Saudi translated both the MSA and the Gulf. The Saudi student was selected to translate two varieties primarily due to MSA being a taught language, i.e., there are no native speakers, thus the pronunciation of MSA is dependent on where the speaker is from. As the participants were to be from Saudi Arabia, specifically central region surrounding Riyadh; a Nijari speaker from Saudi Arabia would be needed for the translations, as they could produce both Gulf Arabic, as the standalone variety, as well as the Gulf-based MSA. This process produced 4 lists of 40 sentences in 4 different language varieties: English, Egyptian Arabic, Gulf Arabic & MSA (4 language varieties x 40 sentences).

4.5.3 Equivalency

The translated sentences were checked for equivalency to the original English sentences, specifically in relation to the accuracy of the sentence. The equivalency checks were conducted by three independent linguists (two Saudi Arabians and one

Egyptian), who were given the original English sentences and the Arabic translations in written form; they were asked to assess whether the sentences' meaning was retained in the Arabic sentences. Any issues identified within the Arabic sentences were relayed to the original translators to fix. All fixed sentences were subject to equivalence checks until the meaning was considered suitability retained by all three linguists. Moving forward only the 120 Arabic sentences (total of the 3 lists) were continued for use in the study (3 Arabic varieties x 40 sentences).

4.5.4 Recording Sentences

The three female speakers were recorded performing the translated sentences. An Egyptian recorded the 40 Egyptian sentences, a Saudi recorded the 40 Gulf, and another Saudi the 40 MSA sentences. All recordings were conducted in a sound-proof language lab; designed for phonetic research and recordings. The speakers were asked to produce the sentences in an even and natural manner; to assist with this, the individual lists were recorded in single sessions (of 40 sentences in a single recording).

There were three recordings for each list, to provide comparisons for speech speed, error and providing backups if there were data failures or losses. The speakers were instructed to pause for at least 1 second between each sentence, in order to replicate a continuous and natural question and answer discourse pattern. The pauses were also to allow for the speech positions of the speaker to reset for each sentence, to avoid the final point of the previous sentence affecting the next.

Overall, the recording sessions produced three recordings of each sentence for each language variety, thus there were a total of 360 recordings (3 Arabic varieties x 40 sentences x 3 recordings). Given that the recording sessions were single events, the researcher used Audacity® 2.0 (Audacity Team, 2008) to extract the audio for each sentence into a separate sound file for use. The extracted sound files were all checked for consistency of onset-time, to prevent some recording having longer time between the commencement of the audio file and the onset of the sentence recorded. This process allowed for the collation of 3 sets of recorded sentences for each language variety. The total of 360 were put forward for checks for clarity and accuracy of variety checks.

4.5.5 Recording Checks

Post-recording checks were conducted, the checks were to assess: the production accuracy of the speaker, the background noise & any sound file damage. The checks were performed by three Arabic-speaking postgraduate linguistics (one Egyptian national fluent in Egyptian, and two Saudi Arabians fluent in Gulf; all three reported being fluent in MSA), who were not involved in the stimulus production process thus far. The speakers were asked to confirm that the language features, such as accent & prosody, were representative of the language variety's speaking population. The Egyptian assessed the 120 recordings of the 40 Egyptian sentences, whereas the two Saudi's assessed the other 240 recordings of the 40 Gulf and MSA sentences. Through this process it was confirmed that the recorded speakers were representative of the target variety, i.e., the Egyptian's language production was identifiable as being native-like for Egyptian Arabic, these checkers were critical, as

they ensured that both the sentences, and the sentence production in recordings, were representative of the target variety, and not representative of another colloquial Arabic variety.

During this process, a single sentence from the Egyptian list was removed. This was due to the recording checker spotting a previously unidentified ambiguity in the lexical choice made in the translation stage. The checker identified that there were two different lexical items that could be used to accurately describe the key word, differing only in formality. Of the two words, the translation contained the less-frequent word, which held the higher formality but was recognised as anomalous in standard usage by the checker, as such, this sentence was deemed unusable for the study. Ideally, at this stage, the sentence in question would be replaced with a new sentence from the original list (see *4.5.1 Sentence Extraction*) and subjected to the same steps as the other sentences, including recording. However, this was impractical to achieve, as the speaker for the other 39 sentences in the Egyptian list was no longer available for recording sessions. In the interest of practicality, the sentence was removed from all of the lists, thus reducing the total number of sentences available for use to 117 (39 sentences x 3 Arabic varieties).

Following the checks, a ranking task was conducted to identify which of the three recordings performed for each sentence was the best for use, the criteria for this was clarity of speech. This task was conducted by the recording checkers, who were asked to rank the 3 recording of each sentence by the criteria as a group. The results of this task produced 39 sentences per Arabic varieties which were rated as clearest

in speech, with a total of 117 recorded sentences for use (3 Arabic varieties x 39 sentences x 1 recording of clearest speech).

4.5.6 Allocating Sentences for Procedure

When presenting these recordings to the listeners; it was decided to not repeat the sentences across the Arabic varieties, to avoid a priming effect, as the use of all 39 sentences in each variety would result in the 2nd and 3rd instance of hearing the sentence to be potentially impacted by the previous instance heard. To counter-act this risk of priming effect, the list of 39 sentences was sorted into three different lists, three groups of 13 sentences. The lists were then allocated to one of three Arabic varieties, whereby the sentences would only be played to the participant in the variety allocated, for instance, sentence 12 was allocated to Gulf Arabic, so the recordings of the sentences in MSA and Egyptian were not used in the study going forward. This produced a play list of the sentences; with the Arabic variety each sentence was to be played in pre-determined.

The sentences were then allocated further for the two stages of the experimentation. 1 sentence was needed for the practise phase, and 36 were required for the main phase (12 sentences x 3 Arabic varieties). For the practise phase, a single sentence was required; to selected which Arabic variety the practise sentence was to be in a random number generator was used, which selected a sentence from the Egyptian list. This sentence was removed from the Egyptian list for selection in the next task, leaving 38 sentences available (comprised of 13 MSA, 13 Gulf & 12 Egyptian sentences). The list for the main phase was to contain an equal number of sentences for each language variety, which was 12. All of the remaining Egyptian sentences

pp. 311

were allocated, and 12 of the 13 sentences in the Gulf and MSA were randomly

allocated. Resulting in two final play list (see Table 18, also see appendix C for final

sentence list).

Table 18: Final allocation of sentences across the Arabic varieties.

	Number of sentences per Arabic Variety		
Phase	Egyptian	Gulf	MSA
Practise (1)	1	0	0
Main (36)	12	12	12

4.5.7 Response Stimuli

For this study, a set of response visual stimulus were produced, based on the

Pearson Test for Reception Of Grammar-2 (TROG-2) (Bishop, 2003). The TROG-2

was designed to measure a person's level of understanding of grammatical

contrasts. The TROG-2 used a picture selection task, where participants are

presented a set of four images, and they have to select the image which accurately

depicts a heard or read sentence. The TROG-2 was selected as the base design

influence to develop stimulus from due to the strict and methodological restriction

criteria for the response images; specifically, that the all the images must be

connected to the heard sentence prior. The four images are variations of the heard

sentence. For example, if an incorrect sentence is selected, it is due to the participant

mis-understanding a grammatical element of the heard sentence, such as a

preposition. Thus, the variation point between the images was rooted in the grammatical units in the heard sentence, such as Gender difference in the Subject (e.g., He/She) or with the relationship between the subject-object (e.g., On the book/ Under the book).

To produce images for the prepared sentences (see *4.5.1 Sentence Extraction*), a criterion for the images was produced. Unlike the TROG-2, the variation points for the response images in this study was expanded from the limit of grammatical function words, allowing for incorrect images based on incorrect parsing of nouns and verbs. This was to account for the different sentence styles in SPIN test; whereas the TROG-2 sentence lists contained similar sentence structures (Bishop, 2003). For this study, a linguistically informed graphic artist was contracted to produce a set of response stimulus image for each of 37 sentences in the final list (see appendix C & table 18). A combination of public domain images, artist original work and photo editing was used to produce the final images. A total of 148 images were produced, with 37 accurately depicting the sentence, and 111 depicting incorrect parses of the sentence (see figure 3 and table 19 for detailed example set of images).
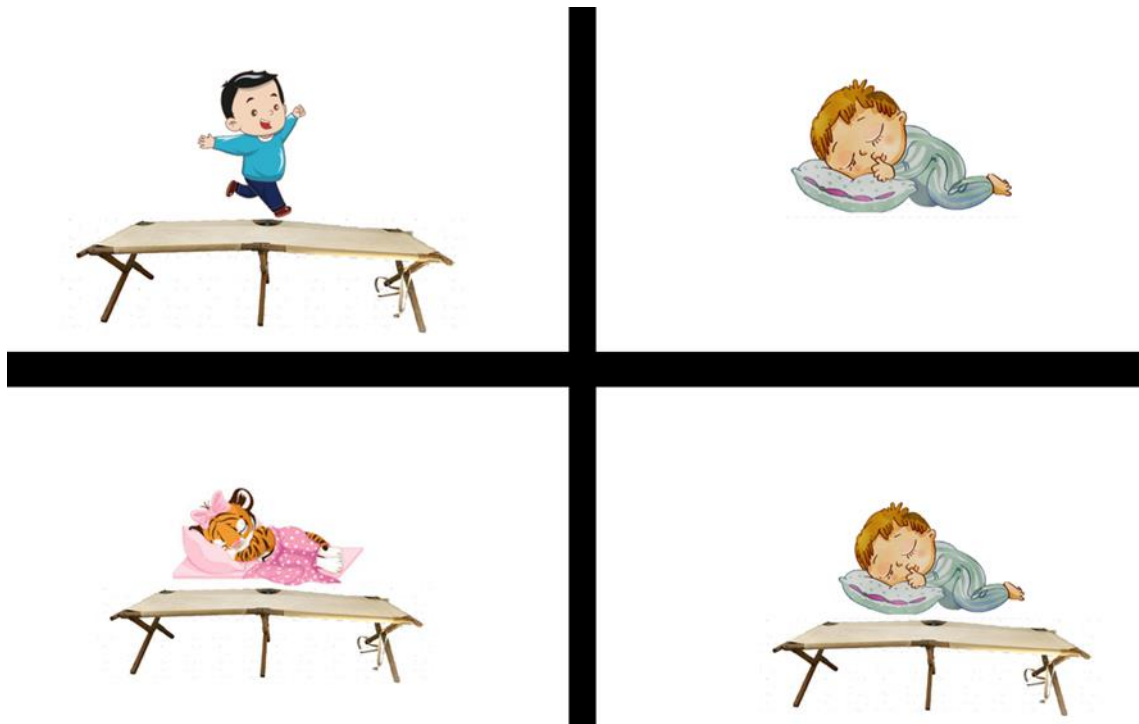
Figure 3: Image options for the heard sentence 'Harry slept on the folding cot'.

Table 19: Image depiction by placement. Variation points from correct sentence are embolden.

Image placement	Depicted sentence
Top-left	Harry **danced** on the folding cot
Top-right	Harry slept on the **pillow**
Bottom-left	**A lion** slept on the folding cot
Bottom-right (correct depiction)	Harry slept on the folding cot

To ensure that the stimulus images were correctly depicting the target sentences, the images were check by 3 independent linguists. These linguists were previously involved with the spoken equivalency checking. They were asked to assess each

image for ease of understanding, with the intention of checking whether they could understand the messaging of the image. To accomplish this, the checkers were asked if they could produce a single sentence describing the image whilst using the majority of the words from the sentence for the correct image. If they were unable, the image was corrected by the graphic artist until each checker agreed that the images correctly identified the intended messaging. Thus, satisfying that each image would present a depiction which was one variation from the heard sentence. In total, there were 148 images, with 4 images allocated for the practise phase; and the other 144 images allocated for the main phase (see table 20).

Table 20: Allocation of images to Arabic varieties sentence lists

Phase	Arabic variety	# of Sentences	# of Images
Practise	Egyptian	1	4
Main	Egyptian	12	48
Main	MSA	12	48
Main	Gulf	12	48

4.6.8 Presenting Images and Stimuli in combination

In both the practise and main phases, the response images would be presented in groups of 4; with the correct image for the heard sentence accompanied with the 3 images depicting variations of the heard sentence. Each quartet would be presented in a crosshair. In order to balance the presentation of the images, the presentation

order and crosshair position was pre-set, to avoid any screen position bias. To decide the order, each section was labelled A-D for clarity (see figure 4), additionally, the images were also labelled A-D (see table 21) allowing for the position of image by screen position to pre-set (see figure 5 & table 22).

Table 21: Labels used for images, based by variation point.

Image Variation point	Label for positioning
No variation	A
Initial noun phrase	B
Final noun phrase	C
Verb phase	D

$$A \mid B$$
$$C \mid D$$
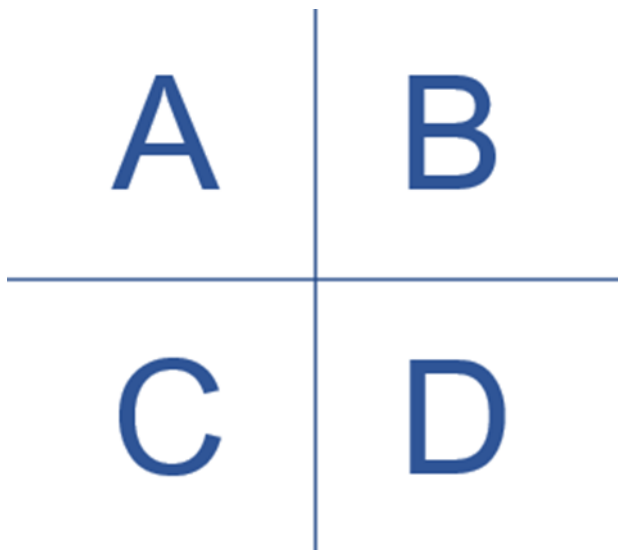
Figure 4: Crosshair screenshot with each quarter labelled for screen position of image
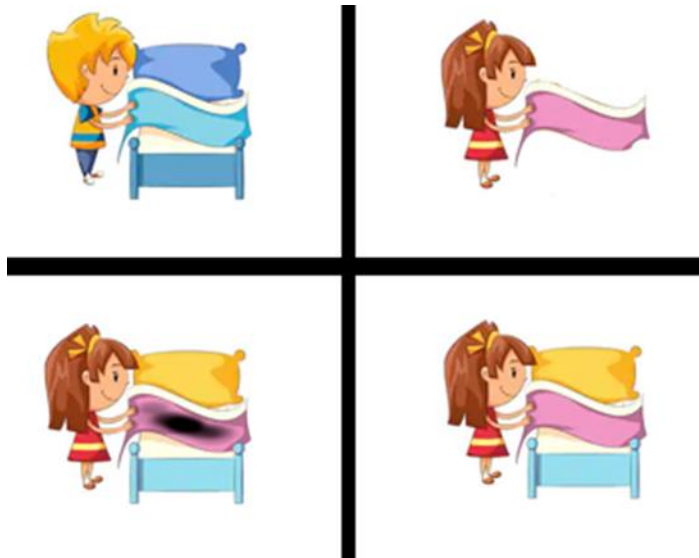
Figure 5: Screenshot of images, when presented in a trial the heard sentence 'She made the bed with clean sheets'

Table 22: Tabulation of images screen placement for image 4.

Sentence List (A-D)	Placement of Image (A-D)	Presented Sentence
A	D	'She made the bed with clean sheets'
B	A	'He made the bed with clean sheets'
C	C	'She made the bed with dirty sheets'
D	B	'She waived the clean sheets'

4.5.9 Allocation of Screen Positions

For the practise phase, the position of the images was matched with corresponding equal label of screen position, i.e., Image A was presented in Position A.

For the main phase, there were 36 sentences and 4 different screen placements; thus, the number of times each image type (A-D) was to appear in each screen position (A-D) was 9. Using a random placement generator; each image position was allocated per sentence until each screen position had 9 occurrences of each image type (A-D). From this, each type of image appeared, across the whole procedure, in each screen position equally, removing any bias due to screen placement (see table 23).

Table 23: Allocation of image type to sentence list & screen position.

Sentences		Number of times image appears in screen position			
Sentence List	# in list	A	B	C	D
A	36	9	9	9	9
B	36	9	9	9	9
C	36	9	9	9	9
D	36	9	9	9	9

4.5.10 Applying Cognitive Load: Stimuli

There were two sets of stimuli developed to facilitate the application of cognitive load on the participants: a set of Mathematical Equations & a set of Colour Blocks. Both tasks were produced in order to inhibit or decrease the working memory, by overloading specific memory functions outlined by Atkinson & Shiffrin (1968) and

Kiyonaga & Egner (2012). Furthermore, the mathematical equations were selected due to the acceptance in psychology that mathematic competence uses a variety of complex mental skills to process and response to both single digit & multi-digit mathematic questions (Raghubar et al. 2010; Saeed & Sasangohar 2017) skills which use the WM. In terms of using different levels of mathematic questions, there is evidence that the demand on the WM varies based on different mathematical operations; for instance, single-digit addition has been found to have greater usage of the central executive than subtraction tasks (Imbo et al., 2007; Imbo & Vandierendonck, 2008). Also, an increased use of WM processing has been linked to operations that use carrying or borrowing operations; particularly there is evidence that the carrying or difficulty of the operation is proportionally linked to the processing requirement in the WM; especially when the operation has to be completed in the mind alone, i.e. without pen & paper (Imbo et al., 2007; Imbo & Vandierendonck, 2008; Klingner et al., 2010).

Whereas the colour blocks utilised visual WM elements; as Todd & Marois (2004) identified colour tasks as demanding of the visual memory processing in experimental conditions, and the colour block also utilised the short-term memory as the blocks were presented in a recognition task (Conway et al, 2005). The combination of both tasks was aimed to facilitate an internal cognitive load, which should overload the Sensory Motor Store, as presented by Lavie (2010). Additionally, the use of two tasks, rather than a single task; is grounded in Engle et al (1999) & Cowan (2008), two projects which suggested that multiple tasks, aimed at different parts of the Sensory Motor Store; would increase demand of the WM and thus

increase the cognitive load. As there were 37 sentences in the language stimuli set (1 in practise, 36 in main task); there were 37 tasks per cognitive load set, which totalled 74 tasks (2 cognitive sets x 37 tasks); this was designed to outnumber the language task to produce the cognitive load effect, as the participant would undertake cognitive and language tasks at a ratio of 2:1.

4.5.10.1 Mathematic Equations

A random equation generator was used to produce the whole set of 36 equations across three difficulty levels: easy, medium, and hard (see table 24) for the main phase. A single easy equation was generated for the practise phase. Easy tasks made use of single digit operations, whereas the medium and hard contained multi-digit operations, with the hard tasks having the added complexity of decimal places. There were more questions of a greater difficulty to facilitate higher WM usage and cognitive load effect (See table 24). The options in the response blocks were variations of the correct answer, for instance the wrong decimal place or numbers in the wrong order (see figure 6).

Table 24: Mathematical operation dataset breakdown with examples

Difficulty	Limit of Operations	# of Operations	Example Operation & Answer
Easy	Single Digit	8	12 - 6 = 6
Medium	Multi-digit (without decimals)	10	44 x 6 = 264
Hard	Multi-digit (with decimals)	18	59 x 1.6 = 94.4
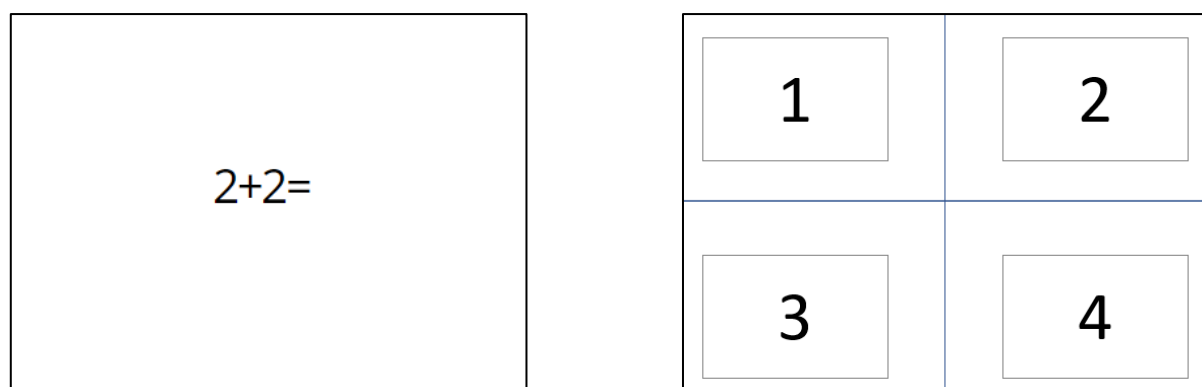
Figure 6: Screenshots of maths task & response options

4.5.10.2 Colour Blocks

A set of 7 core colours were chosen to base the construction of blocks around; core colours were basic colours, like red and blue. The 7 core colours were then modified by saturation levels to produce 4 colour blocks based on the original block colour. Thus, there were 7 colour sets, with 4 levels of saturation variation to produce 28 blocks of colour for use. The saturation levels were regulated by set levels (see table

25), with overall, 7 'Light' colours, 7 'Mid-light', 7 'Mid-dark' & 7 'Dark' blocks of colour, labelled W-Z respectfully, being created. This method was used to expand the options for presentation of the colours in the recognition task as the presented image could have been selected from any of the 28 overall colours in the total set.

Table 25: Colour saturation level per classification group

Category within Core Colour	Light (W)	Mid-Light (X)	Mid-Dark (Y)	Dark (Z)
Saturation Level (%)	0 - 25	26 - 50	51 -75	76 - 100
# of blocks	7	7	7	7

4.5.11 Allocation of presenting stimuli & screen position

The colour task was presented in a similar manner to the language task, in that the responses would be found within a quartile of the screen, as split by a crosshair (see figure 7). As such the same labelling was used (A-D), however the computer randomiser was used to decide the placement of the answer blocks; as this task was to facilitate an effect of cognitive load as a tool, not as the target for investigation. A set list of 39 recognition tasks was produced by two levels of difficulty: Low & High. The low set had responses which had different core colours to choose from when selecting an answer (see table 26); whereas the high set the colours to choose from where all the same core colour, but with differing saturation level (W-Z). Overall, there were 11 answer sets in the low set; and there were 26 in the high set, totalling 37 overall sets for use in the trials without repetitions.

Table 26: Example trial for colour block task

Difficulty	# of trials	Presented Block	Options for Answering			
Low	11	Blue	Blue	Green	Red	Yellow
High	26	Blue	WBlue	XBlue	YBlue	ZBlue

Figure 7: Screenshots of presented colour block & response screen with optional answers

4.6 Procedure

The experimental design was within participants using repeated measures, as each participant was exposed to three different types of stimuli: Egyptian, Modern Standard & Gulf under the same conditions i.e., the language stimulus were audio and were followed by a visual stimuli response task. Additionally, the two cognitive load tasks were conducted in a similar stimuli-response task style. For the cognitive load tasks, the recorded variables were as follows:

Independent – Level of difficulty of the task

Dependent – Accuracy in response (ACC), and the reaction time (RT).

For the language task the variables were as follows:

Independent – Arabic variety exposed in the heard sentence

Dependent – ACC & RT

ACC, which was whether the selected response was correct or not; scored on a binary scale of 1 for correct and 0 for incorrect; & the RT, which was the time between the conclusion of the audio and the selection of an answer, was recorded in milliseconds (ms).

The tasks were run using Gorilla Experiment Builder (2020), an online experimental study service provider (Anwyl-Irnive et al, 2020), allowing the task to be completed anywhere, via phones, computers or tablets. For this experiment, the restriction of platform usage was set as only for touch-screen display devices, such as tablets and phones; to allow for more appropriate comparison of the effects and responses. Additionally, participants were asked to wear headphones & had to confirm they had working sound before the starting of the experiment.

There were two sections to the procedure: a practise and a main. The practise contained a single trial; whereas the main contained 36 trials; with 12 trials per Arabic variety. There was a break between the practise and main, to allow the participant to fix any technological issues & prepare for the main session. Overall, there were 1548 trials completed across the 43 participants; within which contained 516 trials for each of the 3 Arabic varieties. Therefore, there were 516 ACC & RT scores per Arabic varieties & 1548 ACC & RT for the two cognitive load tasks.

There were three tasks overall per trial, also once each task was completed the next began immediately; thus, there was no break in the main session. The order of trials

was randomised using the randomisation function in Gorilla; the tasks within the trials were also randomised via the same function. The only set feature was the order of the trials. The order of tasks within each trial were as follows (see figure 8 for visualisation):

Number Calculation:

The participants were presented a mathematical question for 3000ms. After the time elapsed, another screen was presented with a selection of 4 possible answers for the question. One was correct & the other three incorrect. The placement of the correct answer was randomly generated by Gorilla. The participant was asked to select the correct answer.

Colour Recognition:

The participants were presented with a block of colour around a white background. They were given 3000ms to memorise the specific colour block. Similar to the previous task; after the time elapsed, they were shown a set of 4 options, with one correct & three incorrect. The placement of the correct answers was also randomised by Gorilla. The participants were asked to select the option which contained the exact image shown before.

Sentence Correlation

Participants were presented with a blank white screen whilst a sentence was played. Upon completion of the sentence, a screen appeared with the four images, with one correct, three incorrect in a crosshair design (see the TROG-2 test developed by

Bishop, 2003, for the inspiration of this approach). The participant was asked to select which image correctly corresponded with the sentence heard.

For all three tasks, the reaction time, i.e., the time between the conclusion of the previous stimuli & a response was recorded; and the accuracy i.e. whether the option selected was correct; were recorded. On average, the whole procedure took less than 15 minutes to complete. A full debrief of the aims of the study were detailed upon the completion of the final trial; also, to withdraw the participant could exit the browser tab used for the experiment. There were no withdrawals in this study.
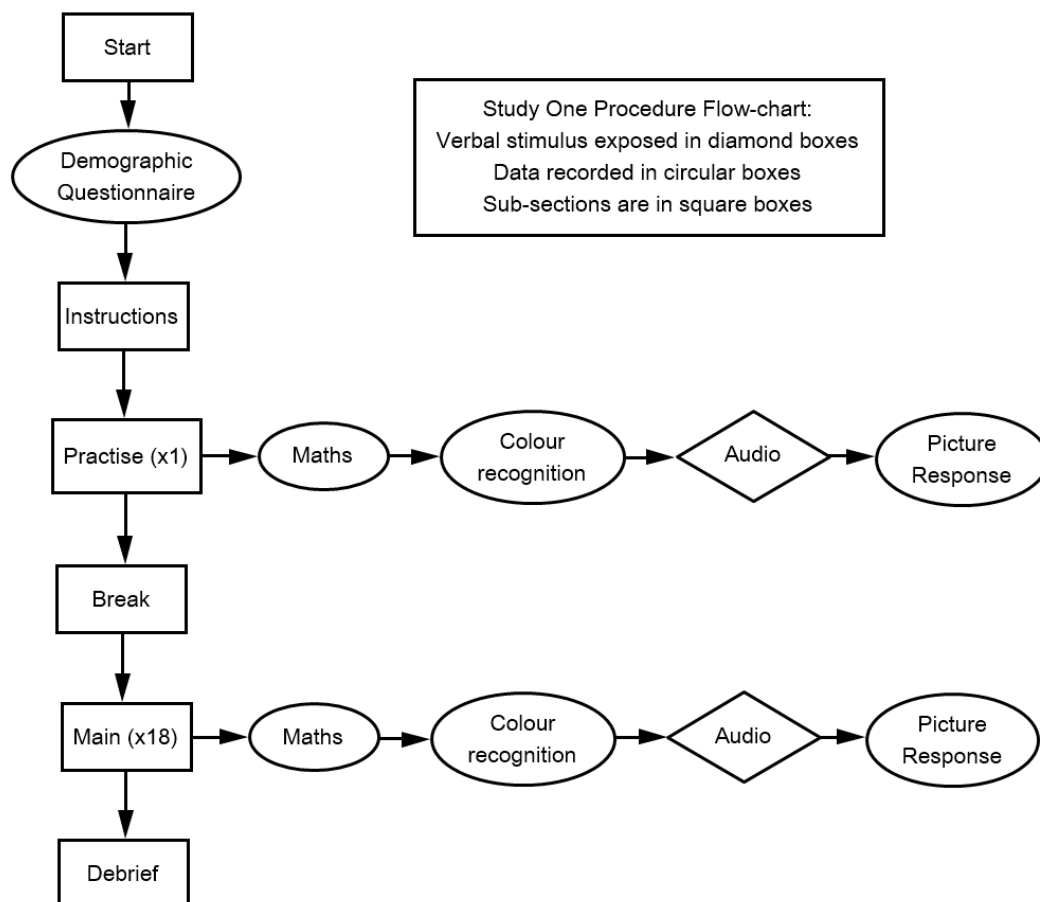
Figure 8: Experiment procedure flow-chart

4.7 Results

4.7.1 Data Processing

The first stage in sorting the data was to calculate the average reaction times and accuracy for each participant for each Arabic variety (the independent variable). This step reduced the number of values (for both RT & ACC) for analysis from 1032 per Arabic variety to 86 (2 values for all 43 participants).

4.7.2 Accuracy Rate

The accuracy rate was calculated using the average mean multiplied by 100, this was due to the accuracy being a binary scoring (categoric). This produced an accuracy as a percentage, rather than a score between 0-1:

$$Per\ particpant\ \left\{\frac{Correct\ Answers}{Number\ of\ sentences} \times 100\ = Accuracy\ Rate\ (\%)\right\}$$

4.7.3 Reaction Times

A single dataset of RT was collated for analysis. This dataset contained the simple mean for all RT recorded in the study per participant when the response given was accurate (see RT average). In the dataset, the times were sported according to which Arabic variety was exposed in the trial, thus the dataset contained the scores split into three groups (Gulf, Egyptian and MSA).

4.7.4 Participant RT average

This dataset was produced by collating all of the RT recorded in total, which was 1548, (36 sentences x 43 participants), then removing all RT values where, in the

trial, the response was inaccurate (scored as 0). A total of 695 values (325 Egyptian, 371 Gulf and 342 MSA) were retained in this reduction, which were then grouped by which Arabic variety was exposed. There was a further reduction by calculating the average RT of all accurate scores per participant for each Arabic variety, which reduced the dataset to containing 129 reaction time values, which was 3 values (one per Arabic variety) per participant (1 RT x 43 participants x 3 Arabic varieties):

$$Per\ participant \left\{ \frac{\Sigma(RT\ when\ Acc = 1)}{Number\ of\ RT\ when\ Acc = 1} = Average\ Accurate\ RT \right\}$$

4.7.5 Testing Groupings

To analyse the whole dataset, first it was investigated whether all the participant's data could be treated as single group. To test this, the following tests were applied for the variables of Sex & Age respectively; to evaluate whether the dataset was homogenic, or if the data should be treated as two different datasets, based on Age or Sex category.

A MANOVA was performed on the Participant RT and ACC values, with Age and Sex positioned as fixed factors: to evaluate if either demographic factor had effect on the dataset. There was no effect identified for Age (p = .431, F = .847, n2 = .013) or Sex (p = .849, F = .164, n2 = .003). From these results, the lack of difference or effect of the demographic factors was used to justify treating all values as a single dataset.

4.7.6 Outliers & Dataset Clearing

A box-plot analysis was conducted in SPSS to identify the significant outliers within the dataset. There were no outliers identified using this technique for both accuracy rates (see figure 9) and reaction times (see figure 10).
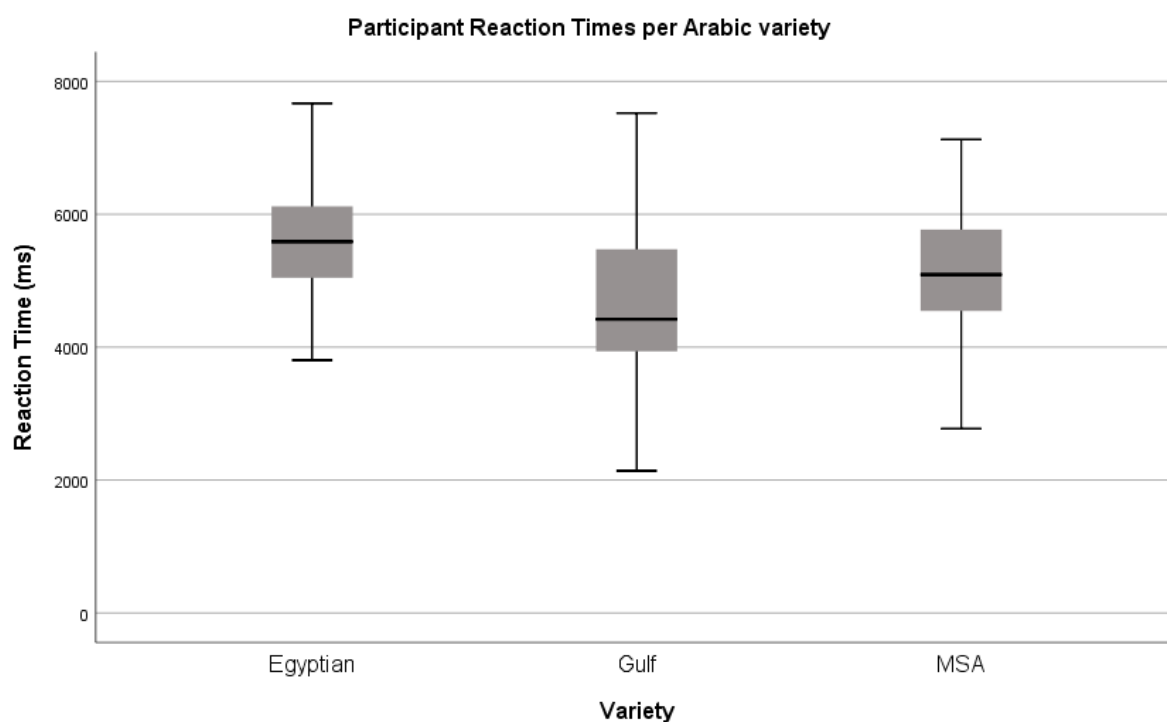
Figure 9: Depicting a Box-plot analysis for participant reaction times grouped by the three Arabic varieties.
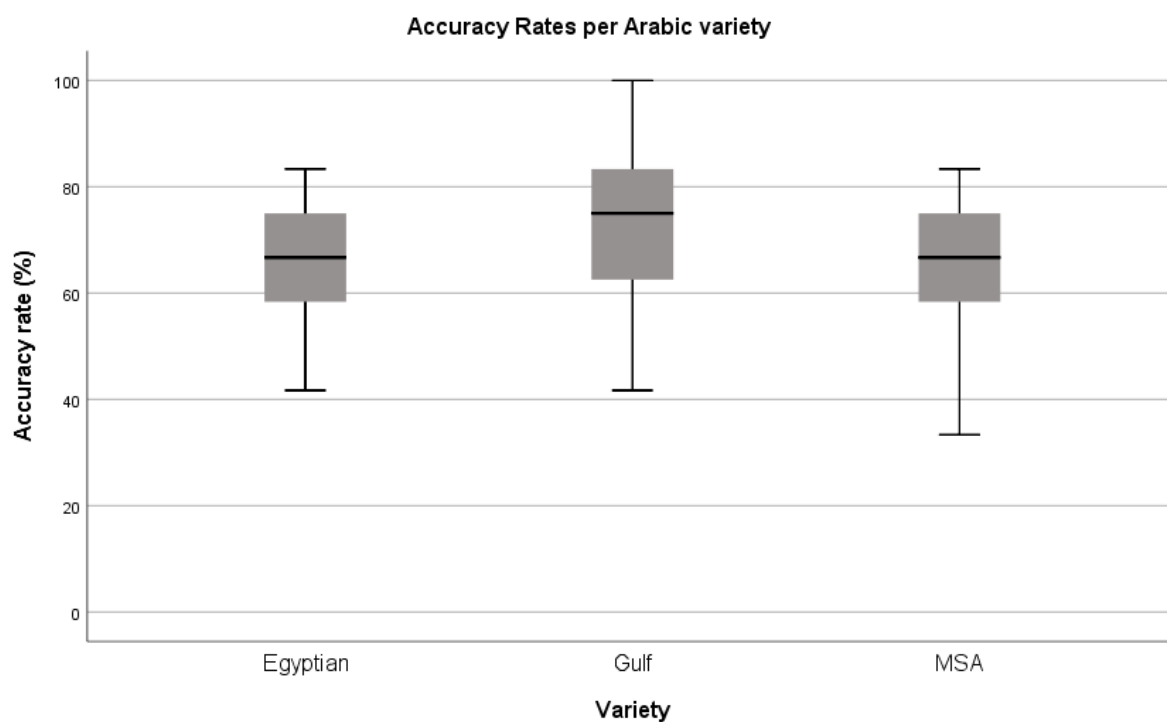
Figure 10: Depicting a Box-plot analysis for participant accurate rates grouped by the three Arabic varieties.

pp. 329

4.7.7 Descriptive Statistics

Following the data-clearing analysis, the cleaned dataset was used to calculate the following descriptive statistics (see table 27) for: accuracy rate, or mean accuracy score across all three conditions (see figure 11) & the RT for the correct ACC responses, or mean reaction times across the three Arabic varieties when the response was accurate (see figure 12).

Table 27: Means & Standard Deviations for the Arabic varieties' groups

Arabic varieties	Reaction Time		Accuracy Rate	
	Mean (ms)	Std. Deviation (ms)	Mean (%)	Std. Deviation (%)
Gulf	4626.34	1095.73	71.90	15.11
Egyptian	5622.74	1044.33	64.34	11.40
Modern Standard	5102.69	966.54	66.28	12.72

4.7.8 Inferential Analysis

This analysis will report on the variables in the following order: Accuracy rate, followed by reaction times of accurate responses.

To investigate the relationship between the accuracy rates and reaction times between the three Arabic varieties, a MANOVA was performed on the dataset, finding a main effect for both accuracy rate (p = .025, F = 3.820, n2 = .057) and reaction times (p < .001, F = 9.933, n2 = .136).

A post-hoc Bonferroni test was conducted to examine the relationship between the three Arabic varieties, for both accuracy rate and reaction times scores, to identify where the differences occurred.

With accuracy rates (see figure 11), the differences were identified one of the relationships; Egyptian – Gulf (p = .026).

Furthermore, with reaction times (see figure 12), the differences likewise located with the Egyptian – Gulf (p > .001) relationship.
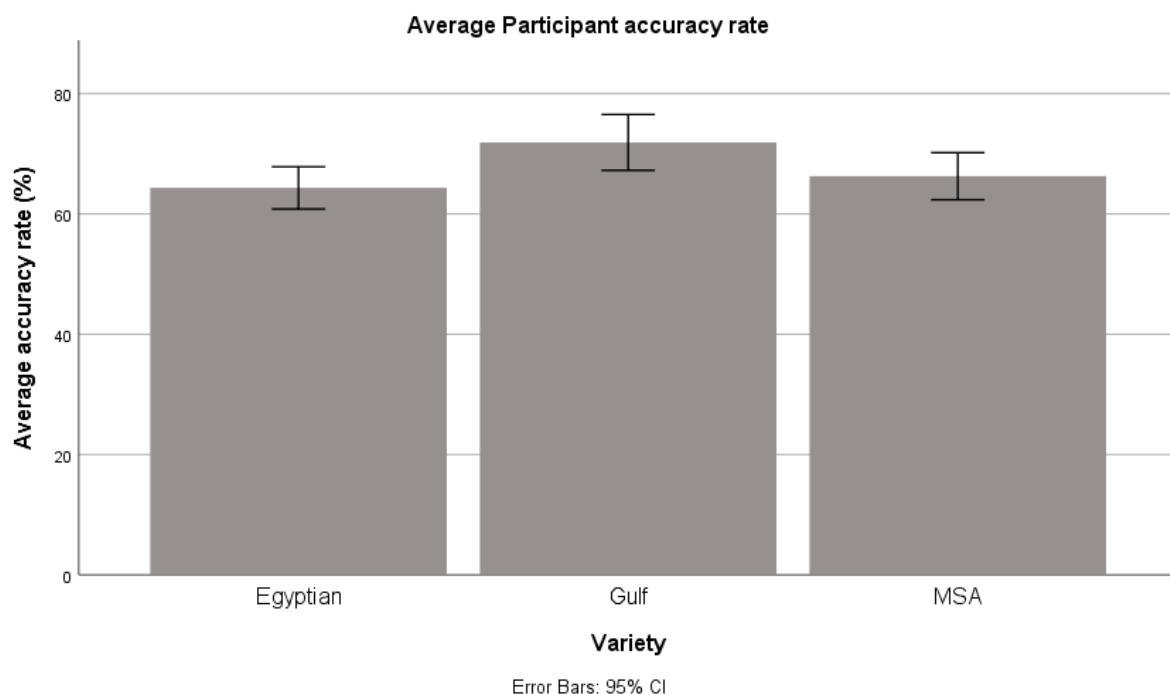
Figure 11: Bar chart of mean accuracy scores by Arabic varieties. Error bars indicate 95% confidence intervals.
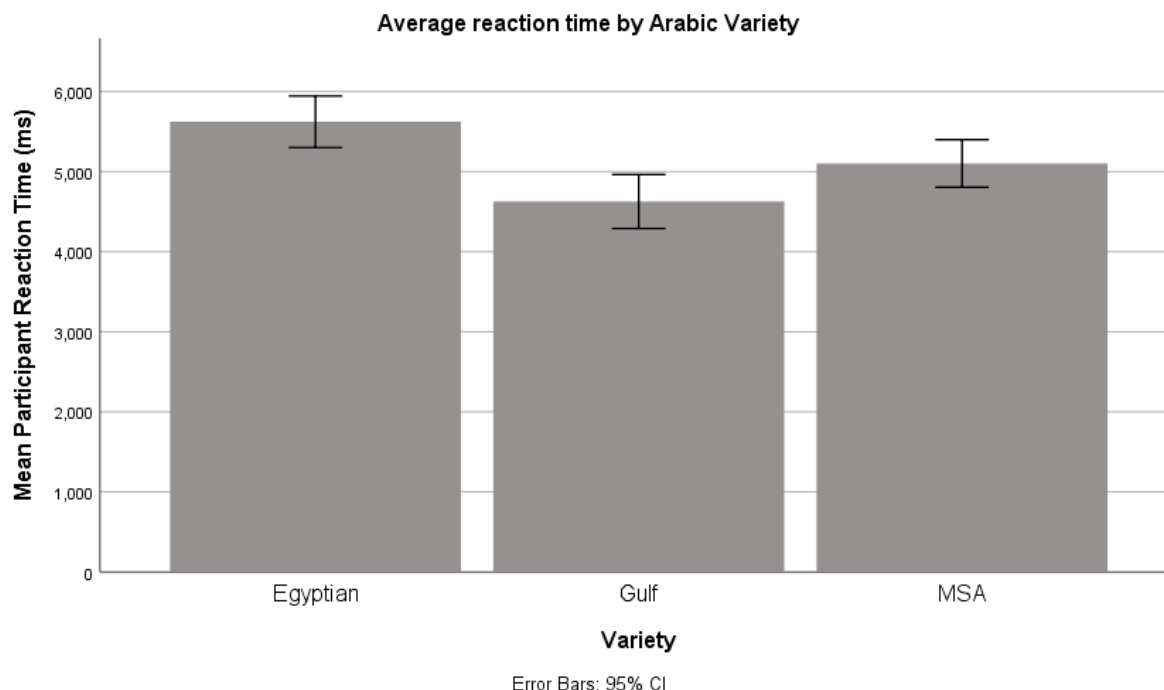
Figure 12: Bar chart of mean reaction times of correct responses by Arabic variety. Error bars indicate 95% confidence intervals.

4.8 Discussion

The results of this study address the three hypotheses presented as well as inform the overarching issue of Ausbaucentrism, and the deployment of Abstand approaches to categorising languages; particularly the issue of thresholding, and the distance required to objectively record and categorise an Ausbau language into multiple language varieties using the Abstand technique of measuring mutual intelligibility. In a similar fashion to the previous chapter, this discussion will be structured to address each of the three hypotheses formulated, in order of appearance. After this, the issue of Abstand thresholding and mutual intelligibility will be discussed, specifically focusing on the usability of intelligibility as a criterion for the demarcation of language, and the extent to which language classification can be standardised using objective measures. The discussion on Abstand thresholding will

be limited to exploring the acceptability of classifying any of the three language varieties tested in the study as being distinct languages, by virtue of data availability and the scope of the study itself.

To remind the reader, the three hypotheses were:

1.   The local variety of QA will be the best result for both accuracy and response speed, in comparison to the other Arabic varieties exposed, (as the local variety is the primary language for the listeners and therefore is the most established and exposed form of Arabic to the participants).

2.   Colloquial Arabic varieties will be responded to with greater accuracy and at a faster speed than standard Arabic, (as colloquial Arabic's share linguistic features but these similarities are not shared between colloquial and standard Arabic's; therefore, the processing will require greater effort which will slow down responses).

3.   Modern standard Arabic will not score in the middle of the other Arabic varieties, in terms of performance in responses (in a rejection of the Ausbau centric idea that standard Arabic is the go-between for communication between all Arabic speakers regardless of their origin or their native colloquial varieties spoken).

4.8.1 Hypothesis 1

The findings of this study support the hypothesis, as the responses to the local colloquial Arabic were on average the fastest whilst also being the highest in accuracy. This indicated that the use of the Nijari Arabic as a local Arabic language in an emergency situation is acceptable, as there is a higher degree of intelligibility than the alternative options presented, which were the international standard Arabic and

the Media-dominated Egyptian Arabic. Furthermore, these results can address, in part, the question of whether the local language variety is an acceptable choice to use in a language policy, despite the lack of standardisation and formal regulation applied to the language use in the wider speaking communities. The result of the study supports the use of local varieties when the criterion for use is decided from an Abstand approach, in this case intelligibility. However, caution should also be taken when stating that local language varieties are always acceptable, given that the intelligibility scores, which in this study were reported through accurate responses to the picture-selection task were only gained from the one language community that were used in this study.

In addition, the significance, or lack thereof, between the responses for Nijari and for MSA is interesting, and was somewhat unexpected from the literature, as the linguistic differences between Nijari, as a Gulf QA, and MSA are well documented (see *2.3.3.1 Linguistic Variation*). If a frequentist approach is taken exclusively, whereby a lack of significance is the singular indicator of similarity (homogeneity), then these results would not support demarking Nijari Arabic as being different to MSA. However, in reality, this strong a statement is not motivated from this study overall, predominantly due to the relationship between MSA and Egyptian being equally as insignificant statistically as the MSA to Nijari relationship; yet the relationship between Nijari and Cairenes was significant; which, by a frequentist approach indicated sufficient difference to distinguish both Nijari and Cairenes as different languages from each other. This complexity reflects the overall theme from the literature; that there is disagreement as to the demarcation and classification of

Arabic's (see *2.3.4.5 Intelligibility in Arabic*); this study equally supports and disputes that classification of GA, MSA and EA as being part of the same language. The deployment of an Abstand approach, in a field-work context (i.e., under a form of stress) does not simplify the issue, rather it continues the discussion between the opposing approaches and models of Arabic classifications onto further research. These results can be inferred to supply a partial explanation for the confliction in terminology used in Zaidan & Collison-Burch (2014); as the exclusive description of EA and GA as separate languages is not clear cut, and it may be this lack of clear definition that provides sufficient weakness in current Abstand approaches to allow for Ausbaucentrism to become de facto. Ausbau approaches model languages against simpler criteria, in that the results of which languages exist, and which are sub-languages is based on categorical data or arbitrary decisions, as opposed to the continuous data dependent upon by Abstand counterpart models.

4.8.2 Hypothesis 2

The results from this study support the rejection of this hypothesis. This rejection is identified from the responses as when the Saudi participants heard sentences in Egyptian Arabic, they responded slower and with a lower accuracy rate than similar sentences in MSA. Therefore, the notion that all QA's will always be faster and more accurate cannot be supported, as the results indicate otherwise directly. Additionally, the responses for both QA's were statistically homogenous in relationship, as identified through a lack of statistical significance, and the grouping of both together was not supported by the statistical testing conducted, thus contrasting to

approaches which consider Gulf Arabic and Egyptian Arabic as being part of the same QA (Horesh & Cotter, 2018).

These results inform the debate on linguistic distance and intelligibility between Gulf Arabic and Egyptian Arabic. Versteegh (2014) and Horesh & Cotter's (2018) defence of Mashriqi Arabic's being considered equal and thus a single group of language was motivated by an assumption that the linguistic differences between said Mashriqi colloquial Arabic's were insignificant in comparison to the similarities shared. The primary examples of similarities included that both share commonalities within phonological systems (see Holes, 1990 for Gulf inventory and Watson, 2002 for Egyptian inventory), as well as the syntactic; given that the QA's share the same word order. From the extent of evidence supporting of the shared features between the two Arabic's, it was expected that this linguistic closeness would allow for high intelligibility between both varieties; a result which was not representative in this study. In fact, the opposite was recorded, leaving in question the impact that shared features have on the intelligibility of another language variety in the context described in TWB (2017a). Studies investigating the distance of the two colloquial Arabic's have focused on the systematic differences between them (see Bassiouney (2009), Akbacak et al., (2011) and Kwaik et al., (2018b) for QA comparisons), thus the contextual use of the language varieties has not been considered. This lack of context might be due to the distance being fixable in conversations in general usage, particularly, by re-phrasing or repetition of unfamiliar features when they cause communication issues, and as such, general Arabic linguists assume that the intelligibility is always at acceptable levels, in that it does not impact the overall

discourse functions i.e., the transfer of information. However, the ability to repair or reassess language use is not contextually available in disasters, as time is sparse, and the deployment of language is limited (see *2.3.1 Current on-the-ground communication strategies used in emergencies*). Therefore, this study's rejection of the intelligibility assumption highlights how contextual factoring is vital for understanding the situation presented in TWB (2017a).

4.8.3 Hypothesis 3

The results from this study reject this hypothesis, as MSA was neither the best nor worst in terms of performance in the intelligibility tasks, therefore supporting MSA being considered an acceptable middle-ground variety for use. Additionally, and most interestingly, was that statistically the results of MSA could be grouped with either GA or EA, as there was insufficient distance in response times and accuracies to classify MSA as heteronomous from either QA; however, the relationship between GA and EA was heteronomous, as such the grouping of all three varieties in one was not supported. This relationship supported the notion that MSA can be used in replacement of a colloquial Arabic, as the results indicate that swapping GA for MSA would not significantly impact the intelligibility overall, as opposed to swapping GA for EA which would result in a reduction of intelligibility.

The fact that MSA could be grouped with either QA supports the ideology that MSA should be used when discussions cannot be facilitated in a shared local QA (Versteegh, 2014). When reviewing the literature, this ideology was categorised as Ausbaucentric, as the evidence provided was anecdotal or qualitative, which is common feature of underlying Ausbaucentrism (Leonardi & Tamburelli, 2021), yet the

results from this study in fact support this assumption for MSA being a go-between. This calls into question whether Ausbaucentrism is as de facto as argued by Leonardi & Tamburelli (2021), particularly the argument that language ideologies and commonly held theories on Abstand measures, such as intelligibility or usability in context, are always produced through Ausbau approaches. However, this conclusion should be considered with care, as the results from this singular study do not provide sufficient evidence to support the opposite, i.e., that Abstand is always considered as Abstand, as that is already established as untrue (see Trudgill, 1992). For a reminder, consider the agreement in Arabic linguistic that intelligibility between Moroccan Arabic and Gulf Arabic is low enough that communication cannot be facilitated between speakers of both varieties (Versteegh, 2014); yet both are still considered the same by political authorities, such as the UN and the AL (Djennane, 2014).

A co-variate factor to consider is the social expectation that MSA is the appropriate Arabic for inter-community dialogues, and thus participants may have had greater confidence when responding to MSA stimulus than to EA. For a reminder, when a person is under-pressure, whether that be socially or cognitively, they will rely on the underlying beliefs to motivate actions, such as measuring the veracity of information received based on the identifiable features and characteristics of the speaker (Shouby, 1951; Chakrani, 2015; Mirshahidi, 2017). From the literature, it was established that Arabic speakers categorised themselves into multiple social groups (Hachiumi, 2013; Bassiouney, 2014) and the shared geo-political Mashriqi identity was one of the recognised groups (Bidaoui, 2016) alongside national identity (Holes,

2011). It could be the case that the shared national identity was the primary group considered by the participants, rather than the pan-regional identity, which would suggest that the messaging in GA and MSA would be viewed more positively than the EA. GA and MSA are often found in daily use in Saudi Arabia, and as such, the use of them is considered a shared social feature of the national identity (Holes, 2011); a grouping which could view EA as being an outsider to the national group. This rejection, if due to social rejection, could explain why the EA responses were significantly slower and lower in accuracy than the GA. When responding to GA, the participants would have considered the speakers as being part of the community, as GA is the local QA used in most domestic discourses; and with that, they would have little reason to suspect that the message was wrong or that their understanding was inaccurate; thus, the responses would be faster, and their underlying intelligibility would show. However, the EA could have been received with scepticism, and thus reassessed by the speaker before deciding on the accurate response. This delay could explain the significant time differences between GA and EA; as the reassessment would have increased the demand on the working memory as the whole message as a unit would be reassessed, meaning the information received would be stored for longer, resulting in weaker processing and information retention rates (Engle et al., 1999; Conway, 2005; Conwan, 2008). The result of MSA as the middle-ground would therefore be due to the linguistic differences being realised and fixed, with an individual who was not reserved in trusting the information received. In essence, the reassessment of the overall message would take longer than the continuous repairing of parsed unfamiliarity's (or linguistic differences) (Li & Yang,

2013); which would indicate that the MSA was fastest due to the underlying belief's supporting trust of messaging over the messaging from an outsider variety. The role of trust between Arabic speakers is an under-researched area (as established in *2.3.5.3 Status quo of Ausbaucentrism and Diglossia in Mashriqi Arabic social attitudes*), yet the impact of social trust, and the belief that MSA is a go-between may have influenced the results of this study.

4.8.4 Overarching issue: Thresholds and mutuality

The results of this study provide insight into the current debate pertaining to operationalising intelligibility as an Abstand demarcation tool for language planning. In practise, the results support the use of intelligibility, as a criterion, for empirically investigating and measuring language use and understanding between multiple language varieties, as shown by the scoring of accuracies recorded and processed in the study. Furthermore, the testing design was not specifically limited to the three Arabic varieties tested, and as such, could be deployed to explore the relationship between other language varieties.

This study's results do not provide clear support on any of the language demarcation thresholds identified from previous literature, nor do the results indicate from themselves a clear demarcation between all three of the Arabic varieties included. From the results, both arguments of language demarcation were supported, those of MSA as a go-between language variety, as that QA's between Egypt and Saudi Arabia are distinct enough to be separate languages. The former point is a key defence used by Ausbaucentric organisations for the rejection of marking colloquial Arabics as separate languages; based on the ideology that the standard Arabic can

still be understood equally across the regions (Amara, 2018; Versteegh, 2014). Whereas the latter is a primary classification supported by Abstand-method studies (Farghaly & Shaalan, 2009; Akbacak et al., 2011; Kwaik 2018b), based on the notion that the linguistic differences are great enough to impair communication and intelligibility between the speakers of both.

The Ausbaucentric argument is supported by the lack of (statistical) difference identified between the individual QA's and MSA, which indicates that the responses to MSA were the same as the responses to both QA's. This situation indicates that MSA should be used in a language policy, as on-balance, the listeners will have adequate understanding and response speeds to messaging; unexpectedly supporting the use of MSA in emergency context, indicating that the support for MSA may not be exclusively due to Ausbaucentric attitudes. However, the relationship between the QA's supports the alternative view, that of splitting Arabic into multiple languages (Zbib et al., 2012), in which case breaking the Ausbaucentric status quo. Whilst the individual relationships between QA's and MSA was homogenous, the comparison between both QA's indicated heterogeneity. This indicates that the use of GA was markedly better than using EA if the population for intervention is from Saudi Arabia, which indicates support for the familiarity effect (Gass, 1984 & Alresanini, 2016), which presents that a localised QA would be the most understood (i.e., intelligible) language variety. Whilst this study did not measure the exposure to the three Arabic varieties directly through the demographic questionnaire, the selection of the varieties was done based on the average exposure from the literature. The localised QA was Gulf Arabic, which is used in domestic and general

communication, Gulf-accented MSA is used in schools and governmental business, whereas Egyptian was selected due to the prevalence of the variety in the media industries, both modern (Hachimi, 2013; Abdul-Mageed, 2018) and historic (Versteegh, 2014). This selection criteria allows for the familiarity to be considered as an assumable feature; in which case the results support that exposure to a language variety may be an underlying cause for higher intelligibility despite there being linguistic differences reported between the variations in question. It is not a nuanced suggestion that exposure to grammatical differences can improve communication (see Gooskens et al's (2015) and Schüppert & Gooskens (2012) work on child intelligibility and exposure), however, the contextual constant in this study (i.e., the cognitive load) implies that this exposure counter-acts linguistic differences in the field. The counteraction can explain the issue of colloquial Arabic's being less understood in the TWB (2017a) report, than the use of MSA. To recap, there were instances where interpreters were having to re-interpret speech between two Arabic speakers, as neither could understand the QA their counterpart was using. Situations, which were unexplainable from the literature, as the QA's were often perceived as mutually intelligible, when in fact, this study identified similar results, thus offering insight into one of the factors that could have caused the situation itself.

Furthermore, the highest accuracy score within this study was found with the varieties of greatest exposure, with around 72%, that of the localised Arabic variety, which could be used as a threshold for mutual intelligibility in future studies. This threshold is based on the notion that localised QA's retain the highest intelligibilities, as well as the defence that the localised QA is representative of the local community,

which functions in a state of mutual intelligibility. Therefore, this study supports

considering the threshold for emergency mutual intelligibility at around 71% or above.

Likewise, the results also provide support for the lower threshold, for when a

language variety can be considered intelligible, but not to the extent of mutuality

between speakers, in which cases the intelligibility would be around 64% or lower.

This lower end is based on the results for EA, which were significantly different to the

high intelligibility GA, and thus cannot be considered as equal. If GA is a measure for

mutual intelligibility, then EA cannot be considered mutual as the results were

heterogenous, as the responses were noticeably weaker in accuracy.

In practise, on a language policy level, these results support Zbid's (2012) efforts to

split Mashriqi Arabic's into two different classifications of language, with Egyptian

distinct from Gulf, as both varieties are responded to differently, as shown in this

study. Furthermore, the evidence of excluding intelligibility in disaster language

policies is weakened, as the results clearly indicate that the Arabic varieties of the

peninsula are not part of a single language which is not impacted by language

variation and distance, as the responses to different varieties were noticeably

different. However, whilst the results also support the consideration of using MSA as

a default Arabic to use in emergency context, it is worth noting that the production of

MSA is reliant on a QA for borrowings of speech conventions such as phonotactic

information and phonological inventories. As such, the prevalence of MSA in this

study could be due to the use of Gulf-accented MSA, whereby the MSA was realised

using the GQA production rules as the basic guidance. The impact of the guiding QA

on MSA production is under researched, which could be due to the Ausbaucentric

bias against exploring the linguistic differences and effects that occur when producing standard Arabics, as it counters the ideology that there is a single, unified standard Arabic.

Overall, these results highlight an established issue in mutual intelligibility, specifically that of absolute intelligibility, which is when communication between two language varieties can be considered immutably certain to always be intelligible, providing external inhibitors, such as background noise (See Clopper & Bradlow (2008), Fiedler et al., (2019) and Ngo et al., (2020), are not present in the discourse. It cannot be assumed that Arabic speakers will all understand one another regardless of their colloquial variety, a statement supported in specific cases by the literature (such as between Maghrebi and Mashriqi Arabic). However, this study indicated support for this assumption to be rejected from use when considering current Arabic groupings, specifically, that all Mashriqi Arabic's are mutually intelligible; as the results support the separation of the group, or at least of two dominant language varieties found within, into multiple distinct languages.

4.8.5 Responding to research question 2

> *(2) Are the Mashriqi colloquial & standard Arabic varieties found within the diglossic continuum of the Kingdom of Saudi Arabia sufficiently intelligible, that it does not matter which variety is used in stressed contexts, such as in emergencies?*

Within this study, the intelligibility rates calculated between two Mashriqi QA's and the standard MSA can provide insight into the acceptability of using these languages

in a disaster context. The Saudi Arabian participants responded with greater accuracy to sentences which were delivered in a Gulf Arabic, and the lowest accuracy was reported when Egyptian sentences were delivered. The contrast in responses to both QA's indicates that intelligibility between these two language varieties is not mutual; in that it is not assumable that the intelligibility rate is sufficient to ensure consistent and accurate communication between speakers of these two QA's. The results can be used to support the separation of Najdi Arabic and Cairene Arabic into two distinct languages when compared against each other. This reclassification of the two QA's indicates that the original classification criteria is insufficient to indicate the intelligibility between the QA's. This study highlights how despite QA's holding similar diglossic status's, the intelligibility of the language varieties is not indicative by which language group the variety is categorised as being part of; a common trait with Ausbaucentric language attitudes and classification. The grouping for the Mashriqi Arabic's is predominantly motivated by shared social history between the nation's holding the QA's as their official language (Hachimi, 2013). Although, the position of this thesis is not to out-right dismiss the notion that there are Arabic varieties which are sufficient for communication in a disaster. This is due to MSA scoring between both of the QA's within the study, both in terms of accuracy rate and response times. The results of the MSA responses were homogeneous with both the GA and the EA responses, indicating that the variation towards responding to MSA sentences was sufficient to be close to both QA's, despite the distance between the QA's being significant.

The ideology for considering MSA as a go-between Arabic was considered as Ausbaucentric from the literature, given that most Arabic scholars defined MSA as such without providing empirical evidence to support the classification of MSA as a universal Arabic with Abstand measurements. This study, which used Abstand measures, supports the consideration of MSA as an actionable middle-ground Arabic which is equally intelligible to both of the other QA's tested. Within the Arabic diglossic continuum, MSA is the standard H variety, which is used in official contexts, such as in business or government, as such MSA is marketed as a default inter-national communication variety of Arabic which can be used to facilitate communication between Arabic's from different regions.

Moving forward, the results of this study support the use of a QA, providing that it is the closest QA to the listening population. This study tested two QA's, one which was the local Arabic, and the other was an outsider Arabic; and found that the outsider was the slowest in response and the least accurate when responded to. However, if there is not a speaker available from the local QA, then this study supports the use of MSA as the default go-between Arabic for use in the field, a view which supports the literature (see *2.3.3.3 Arabic Diglossia*). In summary, the distances are sufficient depending on the QA's involved and the target population for communicating with; but in response to whether the QA of choice matters, this thesis takes the stance that it is an aspect to account for, as the intelligibility on an Arabic is dependent of the listening population.

4.9 Conclusion

This study investigated the intelligibility continuum between three varieties of Arabic found within Mashriqi Arabic, a group of language varieties of Arabic which share historical influences, are geographically close and have been classified as being intelligible between speakers (Trentman & Shiri, 2020), regardless of the variation supported or realisation of differences in discourses (Al-wash, 2016). The extent to which this statement of mutual intelligibility, of sorts, is the main topic area for this discussion, as the overarching research aim is to explore the relationship between neighbouring language varieties in relation to the communication framework and the status quo of intra-variety communication. The results of this study indicate support for the distinguishing between Egyptian Arabic and Gulf Arabic, whereby both should be considered separate languages as they are perceived and responded to significantly differently by speakers from Saudi Arabia; however, the results also indicate the intelligibility between Modern Standard Arabic and colloquial Arabic's is not as clear-cut as some historically have referred to (see *2.3.1.4 Abstand and Ausbau*).

Chapter 5: Language Attitudes in Mashriqi Arabic speaking communities

This chapter contains a study which investigated the underlying language attitudes held by Saudi Arabians to speech produced in three different Arabic varieties; Nijari, (as a Gulf QA), Cairene Arabic, (as an Egyptian QA) and MSA (in context) which replicate, in part, the demands of the field describe in TWB (2017a), i.e., under pressure (in this case extraneous cognitive load). This study explored the relationship between speaker trustworthiness and Arabic varieties, specifically whether the messaging itself is viewed differently based on the diglossic status of the Arabic variety used. The aims of this study were to explore whether the negative attitudes between Arabic speakers identified in TWB (2017a) are representative of a wider issue in Arabic sociolinguistics; that of diglossic social prestige interacting with language identification in high-pressure discourses. Trustworthiness, as a feature, was measured through a series of rating scale tasks, with the participants asked to rate five variables linked to trustworthiness, through sliding scales; this produced a continuous variable allowing for the objective calculation of trust; thus, satisfying an Abstand criteria; as results quantified the social distance, or one part of, between the three language varieties of Arabic. Within this study, intelligibility was also tested (in a similar design to the previous study (see *4.5 Stimulus Design*), allowing for investigation into two Abstand measures, both individually and through comparison, to explore possible connections between both, i.e., is low intelligibility indicative, or connected to, low trustworthiness.

5.1 Research Question

This chapter will explore the third research question proposed form the literature review: is underlying distrust between Arabic speaking communities a factor to account for in disaster communication planning?

5.2 Aims

This study aims:

- To address whether the situation reported in TWB (2017a) is representative of an overall issue in Arabic sociolinguistics, specifically, the negative relationships between, and within, Arab societies are realised in negative language attitudes.
- To investigate the status quo of underlying language attitudes, specifically trust, found in Saudi Arabians in relation to speakers of MSA, and two colloquial Arabics used in the nation.
- To identify whether the diglossic status of the Arabic variety impacts the trustworthiness of a speakers by virtue of speech alone, namely, whether the prestige of MSA is retained in practise.

5.3 Hypotheses

In this study there are three hypotheses, motivated from the state of the art.

1.   The speakers using the localised QA will be viewed as the most trustworthy, as the associated attributes support the identification of the speaker as part of the In-group socially

2.   The trusting responses towards MSA speakers will be more positively scored than towards non-localised QA's, as the established social grouping supports viewing MSA speakers as within the community, and other QA's as outside.

3.   The viewed trustworthiness results between Arabic varieties, in terms of performance, will mirror the intelligibility relationship, i.e., the most intelligible will be the most trusted and vice versa, particularly for low intelligibility, for if communication cannot be understood, the veracity of the message equally cannot be assessed, therefore, trust will be low.

5.4 Participants

For this study, there were 60 online participants recruited (M= 30, F=30). Prospective participants were allowed to complete the study if they satisfied the following three filtering criteria:

1)   Over 18 years of age

2)   Has lived in Saudi Arabia throughout their childhood

3)   Is living currently in the Kingdom of Saudi Arabia or have been out of the nation less than one year.

In style with the previous study (see *4.4 Participants*), the participants were recruited via the investigator's personal social networks. The advertisement requested the specific involvement of Saudi Arabian's only, to dissuade other Arabic speakers from attempting to partake in the study. There were two waves of recruitment for this study.

5.4.1 Pre-testing Questionnaire

The participants were screened using a pre-test questionnaire, and in which they were presented a set of possible answers. The answers to the pre-test questionnaire determined whether the participant was accepted for use in the study, using the filtering criteria, whereby selection of answers which did not satisfy the criteria resulted in rejection, such as selecting Egypt as the country of origin with no experience of living in Saudi Arabia would result in rejection.

The questionnaire asked for information related to the following factors: Age, Location, National Identity & Sex. The responses to the questionnaire were limited to a set of pre-decided options, which were presented from a drop-down menu.

In the first recruitment wave the following options were presented (the slash (/) separates the options for reader ease)

Age – 18-24/25-32/33-40/41-50/51-60/61-70/71+

Location - Egypt/Lebanon/Saudi Arabia/Tunisia/Other

National Identity – Egyptian/Lebanese/Saudi Arabian/Tunisian/Other

Sex – Male/Female/Other

From these choices, the accepted answers for continuation were limited for location (with only Saudi Arabia accepted), National Identity (with only Saudi Arabian accepted) and Sex (with only Male &Female accepted). There was no limit of age group during this wave. The resulting participants from this wave favoured (as in all except one) two ages groups (18-24 & 25-32), as such a second wave was

conducted, to narrow the criteria to capture a balanced set of participants. The single participant who selected 51-60 was removed from the study data.

In the second recruitment wave, the selection process for participants was conducted in the same manner, with the only change being the qualifying criteria of age, which was limited down, so that successful participants were only those who selected 18-24 or 25-32. The goal for the second wave was to only recruit the two ages until there was an even balance of ages and sexes in the study.

This goal was achieved as the set of participants was balanced for both age and for sex. With 30 identifying as between 18-24, and 30 identifying as 25-32 and the split of sex was equal between the two age groups (see table 28). At this stage, it was decided that each variable for grouping (both age and sex) would be tested statistically, prior to testing for effects. This group testing would counter-act the recording of ages as categorical data intervals, rather than continuous; as the results would provide justification for the overall grouping of all participants, or the separation of groups, either way, providing justification for all groups to be considered homogenous for statistical testing.

Table 28: Age and sex grouping of the participants.

	18-24	25-32
Male	15	15
Female	15	15

5.5 Stimulus Design

5.5.1 Approach

This study used a matched-guise attitude solicitation design, with cognitive load conditioning. The study was designed initially in British English which was translated into three Arabic varieties (Egyptian QA, Saudi QA and MSA), so that the testing was conducted through the medium of Arabic. Furthermore, the study was comprised of two sections: stimulus response and attitude judgement. For all of the tasks, the stimulus was prepared bespoke. The stimulus response sections contained two tasks: one numeric and one picture selection, which was then followed by the attitude judgement ranking task. The types of stimuli presented were predominantly visual, with the only exception being auditory stimulus, which was used for the picture selection task. The attitude judgement tasks involved evaluating the stimulus presented for the picture selection task. The numeric tasks were mirrored from previous psychological studies (see Imbo & Vandierendonck, 2008 and Hartwright et al., 2018), and the attitude judgement ranking task was replicated from previous socio-linguistic studies (see Schüppert et al., 2015; Nijjari et al., 2019; Abbas et al., 2020) with suitable adjustment of the key words to ensure the desired attitude was measured. The picture selection task was inspired from a pre-existing testing system, the International English Language Testing System (IELTS). The following section will detail the IETLS, and highlight the elements replicated for this study; after which the proceeding sections will outline the criteria for each set of stimuli as well as the steps taken during creation and for quality control of the stimuli.

5.5.2 The International English Language Testing System

The IELTS is a standardised test system for assessing the English ability and proficiency of English speakers. The IELTS is used as the global standard for certifying acceptable English proficiency for employment and educational visa applications (Hyatt, 2013). The test is taken by non-native learners of English to provide accreditation of language proficiency, which can allow for access to English medium courses, such as university degree programs (ibid). The IELTS is used globally across 140 countries to allow access to education and immigration, and the exams are taken by more than five million people per year (British Council, 2022). At time of writing this thesis, the IELTS has been taken by over one hundred million individuals since its inception in 1989 (British Council, 2022).

As a system of assessment, the IELTS is regularly updated and continuously assessed across research topics, from marking pedagogy of English as an Addition Language to comprehension of accent variation in the workplace (Dong et al., 2022). Due to the widespread usage of the test, and the continuous assessment of the system itself, the IELTS was chosen as the foundation for the experimental stimulus; with parts of the IELTS being extracted and refined for the purpose of the study. For this study, inspiration was taken from the language comprehension listening tasks, which are conducted as part of the listening proficiency parts of the IELTS; the specific tasks emulated were the verbal map-based reasoning and direction comprehension, with corresponding questioning strategies.

5.5.3 IELTS Verbal Reasoning and Comprehension Question Style

In the IELTS listening comprehension tests, an examinee listens to an audio clip of an English-medium discourse whilst also answering questions related to the topics and situations discussed by the speakers in the recordings. These questions are designed to assess an individual ability to comprehend and understand continuous spoken English. Within the listening test examples and past-papers, there is variation in the style of the topics; for instance, one question architype is a fill-in-the-blank style, where the question is presented in a sentence, but single words are omitted, and the correct answer needs to be written in by the examinee; in these questions the discourse heard is conversational with multiple participants. Another example is the map-based comprehension, where the question architype is label-on-the-map, where the examinee needs to label the locations on the map, based off the information heard in the discourse (which are directional, rather than conversational). In the map-based testing, the identifiable parts on the map are used as the reference points for the discourses and the question, whereas for fill-in-missing-word tests, the reference points are the words surrounding the blank-space. For instance, in the conversation discourses, the examinee can use the semantic field of the surrounding words as a reference of whether their answer is acceptable, whereas, in the direction discourses, the examinee can use the identifiable features, such as trees, to support their answering. In emergency communication, directions and instructions are widely used as a mechanism to inform the populations of risk as well as providing a method to escape the situation safely. As such, the use of IELTS map-based comprehension

was used as inspiration for use in this study, as the test replicates part of common emergency communication, that of directions.

Each testing series of the IELTS requires the creation of new source material and given that the system has been testing for over 30 years, there is a high volume of past-paper examples available online; all of which could be re-used for testing. However, due to the prevalence of the IELTS within Saudi Arabia, there was a risk that using a previous test's map material could have been presented to individuals who either were tested with the material or have practiced the material in preparation for their own IELTS examination. Therefore, it was warranted to create a new map for experimentation. The process of creating an IELTS style visual map, as well as the discourse statements (directions) was conducted in three steps. Step one involved analysing past papers and study material, with the aim of identifying the basic criteria for both the maps and the directions in the IELTS style; in addition, research from emergency communications was also considered (see *2.3.1.7 Guidelines for emergency communications*), to ensure that the criteria for the spoken directions was reflective of the current practices and techniques of real-world emergency communications. Step two involved creating stimulus using the criteria identified (for both the map and the directions). Step three involved assessing and checking that the stimulus produced satisfied the minimum criteria.

5.5.4 Stimulus criteria

5.5.4.1 Visual – Map

This stage involved the creation of a map for use, the first task was a review of past IELTS papers and practise material. Resources included in the analysis were taken from established providers or assessors of the IELTS, such as The British Council (British Council, 2022) and the IELTS (ielts, 2022), as well as from free-lance IELTS education services, such as Mini-ielts (mini-ielts, 2022). From these repositories, all maps identifiable for listening comprehension were examined, with key visual features recorded, which were categorised into the linguistic and the non-linguistic. Linguistic features included road names and language-based signs, whereas the non-linguistics were identifiable features that were depicted using images. As this study aimed to use three Arabic varieties in the final version for testing, the decision was taken to not include any linguistic features in the criteria, as to avoid any issue of language influence or bias. Furthermore, if linguistic stimulus was used, then the map would have to be completed in all of the tested languages, which is impractical resource wise, and also could result in further issues of bias. Thus, for simplification and rigor, only non-linguistic features were retained as the criteria for the map in this study. A total of fifteen maps were analysed, which produced the following criteria list for the creation of an IETLS style map:

-     Aerial (birds-eye) viewpoint

-     Cardinal point of North

-     Roadways

- Road junctions, such as T-junctions

- Railway line, and junction

- Varies building shapes

- 2-Dimensional

- Minimal imagery detail

5.5.4.2 Auditory – Directions

For the review of the directions given in the IELTS past papers and learning materials, the spoken discourses which accompanied the fifteen maps selected for the previous review were analysed. The following criteria was identified from the previous examples, which were:

- Maximum sentences length of 25 seconds.

- Use of cardinal points for navigation

- No use of grammatical negation (such as do not)

- Imperative form

- Regular use of prepositions (such as above, next to and opposite

- Direction-signpost semantic information pattern relationship (see table 29)

Table 29: Example simplified IELTS direction sentence, separated by the direction-signpost relationship.

Direction 1	Walk west
Signpost 1	towards the blue building
Direction 2	then walk south
Signpost 2	past the green building
Direction 3	then walk north
Signpost 3	towards the brown building
Direction 4	turn north
Signpost 4	at the crossroad
Direction 5	then turn south

Additionally, a second set of criteria was also applied to the sentence stimulus; based off the requirements of current emergency communication, as well as the recommended style and structure of evacuation orders (see *2.3.1.7 Guidelines for emergency communications*). The final criteria list for sentence stimulus was as follows:

- Sentence length, when spoken should not exceed 20 seconds.

- Regular and standardised structure to allow for repeatability, using regular prepositions in a direction-signpost semantic pattern

- Sense of urgency and formality

- Imperative tone

- Culturally neutral

- Length of each sentence being between 20 and 25 words.

- Use of compass orientation for navigation

- No use of humour, subjective terms, discourse markers, abstract concepts (metaphors), grammatical negation or technical jargon

5.5.5 Stimulus Creation

5.5.5.1 Visual Stimulus – Map

A graphic artist was instructed to produce a map that satisfied all criteria identified through the review of IELTS material. During the production of the map, an issue related to the buildings on the map; the artist observed that the requirement of drawing 2-D building shapes in a standardised pattern would result in minimal variation, as architecturally buildings vary little. To counter this, it was decided that the building would be replaced with geometric shapes, which were of similar size but varied by sides and shape: in total 19 shapes were used in replacement of standard buildings. The colour scheme of the buildings was regular, with no deviation; thus, the only identifiable difference would be based on the shape itself. Following this, the map for testing was produced, which included a start point indicated with a red X (see figure 13).
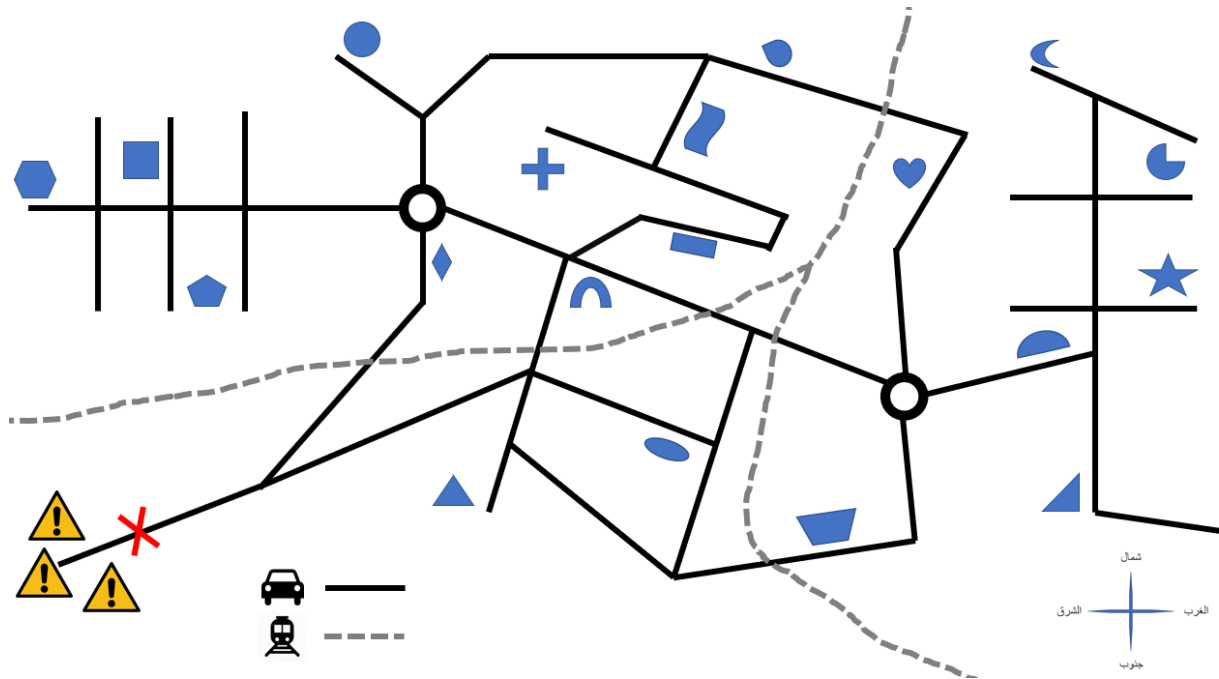
Figure 13: Direction-based map for testing.

The map was compared against the minimum criteria, by the principal investigator, and a volunteer research assistance; there were no issues identified through the check. If elements of the criteria where missed, then the graphic artist would have been asked to fix the issue, then the map would be reassessed; a process that would continue until both assessing parties agreed the criteria was satisfied.

5.5.5.2 Auditory Stimulus – Sentences

Overall, a list of 19 sentences was produced. These sentences were built using the criteria list for sentence stimulus as a guideline. Each sentence contained directions towards a single shape; there was a set of directions produced for each shape; so, there was no repetition of target shape (see table 30 for examples, and Appendix D for full sentence list). Each sentence had at least 4 signposts and 5 directions. The average sentence length was 23.68 words. The number of sentences with 4 signposts was 9, as was the number with 5 signposts. The initial choice (left or right)

was evenly split 9/9; additionally, the term to define the first road choice was also varied equally between the two terms: fork & split.

Table 30: Produced sentences, ordered by sentence number.

Target Shape	Sentence
Hexagon	go left at the fork turn west at the roundabout continue past the junction head west past the square then turn north
Irregular Quadrilateral	go left at the fork turn east at the roundabout continue east at the crossroad turn north at the roundabout then turn south at the junction
Plus-Sign	go right at the fork turn left at the junction go straight over the junction continue around the corner past the junction
Oval	go left at the split walk right at the roundabout turn south at the roundabout head north after the railway turn left at the junction
Crescent	go left at the split turn right at the roundabout turn left past the semicircle continue to the roads end turn left
Heart	go left at the fork head over the railway go north at the roundabout turn right at the junction continue past the railway

5.5.6 Criteria satisfaction

Two checks were conducted on the sentence list. The first check assessed whether the directions themselves were correct; with a clearly defined route and target building. A group of four informants were recruited to check the instructions. They

were each asked to read the sentences and mark on the map the route described, highlighting the signposts. These routes were then compared by the principal investigator, to see if there were any deviations between the informants' markings and the intended direction from the sentences. There were no deviations reported, thus the instructions were accepted as correct and suitably clear to understand. The second check was to assess whether any of the target criteria were violated; this was conducted by the same informants of the first check. They were asked to compare the sentence features against the criteria list for sentence stimulus and report any deviation or violations; zero issues were detected; thus, the sentences were classified as finished.

5.5.7 Translation

This task involved the translation of the 19 English sentences into the three desired Arabic varieties. Two speakers were recruited for this, an Egyptian speaker, and a Saudi (Gulf) speaker. The Egyptian was asked to translate into EA, whereas the Saudi was asked to translate into colloquial GA and the formal MSA used in Saudi Arabia. This resulted in all 19 sentences being available in four language varieties, English & the three Arabic's.

5.5.8 Equivalency

Post-translation analysis was conducted to ensure that the meaning of the translations was accurate to the original sentences in English. This was conducted by presenting the sentences to four speakers of Egyptian (2) and Gulf (2) Arabic; the Egyptians were asked to score the EA translations whereas the Saudi was asked to

score the GA & MSA translations. The scoring was a simple pass or fail, a failure being when the sentence did not accurately relay the full original sentence meaning. If a translation was rated a failure, it would be re-translated until the accuracy was at a pass level. From this stage, there were no issues identified with the translation, and as such the three sentence lists were classified as acceptable for use.

5.5.9 Recording Sentences

Once all 19 sentences were cleared for semantic accuracy in translation, the sentences were presented to a new group of female Arabic speaking informants with the request for the research informants to record themselves speaking the sentences. For the matched guise, there were three speakers recruited for each variant of Arabic, as such there were 9 research informants for this task (3 for Egypt, 3 for Gulf and 3 for MSA). Each informant was asked to record two of the sentences in the list, with one Egyptian speaker asked to record three, this was to reduce the burden on the informants. Each informant was asked to pause between the sentences to allow for the resetting of the mouth positions and to prevent cross-over phonological features and effects, such as pre-voicing, to occur. Each sentence was then extracted from the overall sound file and labelled accordingly.

5.5.10 Recording Checks

To check that the recordings were accurately verbatim of the target sentences, equivalency tests were conducted. The same set of informants from the previous equivalency check were asked to assess whether the speakers accurately performed the sentences, without errors or variation from the intended and expected language

variety. For recap, the MSA of a region used the local QA as guidance for speech production, as such, the checks assessed whether the speaker used Gulf-accented, or Gulf-based MSA terms and convention in their speech production. The informants were also asked to check if there were any anomalous elements, such as vocal stutters or discourse markers in the recordings. In the event that there were any anomalies, the sentences were re-recorded and reanalysed until all informants were content that the sentence was accurate and clear. From this task, there was a single sentence rejected, one from the Egyptian list; as such, the original speaker was asked to re-record themselves. This re-recording was then tested and passed satisfactorily.

5.5.11 Attitude solicitation – Trust scoring

5.5.11.1 Statements for rating

The underlying language attitudes of the participants towards the speakers were recorded using a sliding scale questions strategy, similar to Likert scales (Schüppert et al., 2015; Levon et al., 2021) but without the fixed categories, and the score was a numeric value based on the final position selected on the scale. A set of five scaling statements were created from a single sociological theme: SOLIDARITY. This theme was selected as suitable as the study aim was to measure the status of trust in language attitudes, and solidarity is an established feature of trust in sociological research (Mirshahidi, 2017; Bidaoui, 2020). Each scaling statement addressed a different element within the theme of solidarity, using a different key word; there was little variation between the statements; with carrier phrase regularity being prioritised

to reduce co-variate effects. The key words were based off the personality traits within Solidarity (Abbas et al. 2020) extracted from across the literature, they were as follows: Trustworthy (Mirshahidi, 2017)/ Reliable (Abbas et al, 2020; Bidaoui, 2020)/ Clear (Schoel et al, 2013)/ Honest (Mirshahidi, 2017)/ Respectable (Bidaoui, 2020). Of the five statements, four were presented in an identical carrier phrase, whereas one was adapted to maintain grammatical correctness with the key word. The set of scaling statements were as follows:

This speaker is **trustworthy**

The speaker is **reliable**

The speaker spoke **clearly**

The speaker is **honest**

The speaker is **respectable**

5.5.11.2 Translation and equivalency

To maintain the consistency of using Arabic within the study, the five statements were translated into MSA. This translation was conducted by a Saudi Arabian linguistics student, who was instructed to keep the carrier phrase as consistent as possible. Following this, the translated statements were checked for equivalency by the same informants who completed the previous equivalency checks. The informants were asked to confirm two criteria were stratified: A) the carrier phrase was as consistent as possible and B) that the statements were accurate MSA for the phrase. There were no issues identified in the statements from this check, as such the translated statements were used in the experimentation. For the reader's ease

the rating statements will be presented in English for the remained of the thesis, however, the experimentation itself used the Arabic versions instead.

5.5.12 Scale design

For this study, a 100-point sliding scale was used. This allowed for the participants to select between disagree and agree across 100 intervals. The alternative scale considered was a Likert scale, which presents a smaller set of answers, typically a 5-point or 7-point scale (Kemper et al., 2020). The line between the polar points is thus limited by the categorising of the scale itself, as the maximum number of options is limited by the length of the scale itself. However, the deployment of a sliding 100-point scale provided an expanded range of for selections, as the participant could selected anywhere on the line and provide a viable response. Furthermore, the scores recorded in the midpoint region, which was 10% either way from 50% were considered neutral responses; also, the scores recorded on either side of the scale, which is from the end point and inwards by 40%, were considered supportive of the value at the end point (see figure 14 for an example scale and classification for responses):
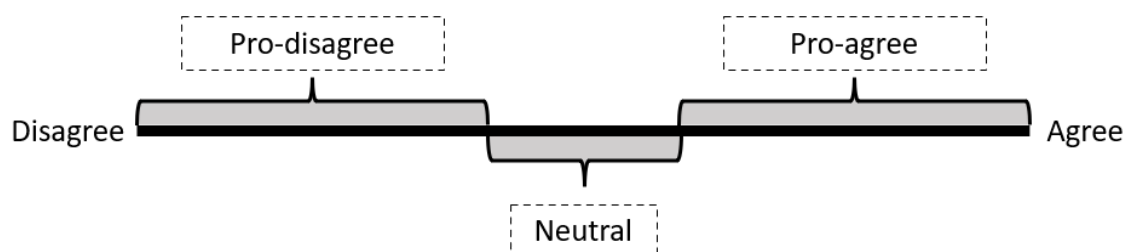

Figure 14: Example scale and classification for responses

5.5.13 Applying Cognitive Load

A set of 19 mathematic equations were produced using a random equation generator. These equations were designed to facilitate cognitive load, and thus to retain a consistent effect a set criterion for the equations was produced, which were:

- Simple operations only (addition, subtraction, multiplication, and division)

- The numerical values for calculation limited to being double-digit at maximum

- Only integers allowed in equation and potential answers.

The correct answer to each equation was collated into a set; additionally, three sets of incorrect answers were also produced for each equation (see table 31 for an example). The incorrect sets were designed to have values that were related to the equation; with two set containing values that would be correct if the operation was incorrectly calculated, such as multiplying by two, when the equation stated multiplication by four. The additional incorrect set contained the reverse of the correct answer, with the digits in the opposite position to the correct list. These sets were used to add additional cognitive load during answer selection, as each option to select was related to either the equation or the correct answer.

Table 31: Example mathematical equation and optional answers

Equation	3 X 15			
Answers	Correct	Incorrect (count)	Incorrect (position)	Incorrect (count)
	45	30	54	15

5.6 Procedure

The main study was Verbal-Matched Guide & Attitude solicitation experiment. Prior to the experimentation stage, a demographic questionnaire was presented to collect consent, demographic information, such as sex; and instructions to complete the experiment. Additionally, participants were asked to wear headphones & had to confirm they had working sound before starting the experiment.

The study was structured into 19 trials, each trial contained a set of three tasks (see figure 15 for visualisation of procedure). The presentation order of the tasks within each trial was pre-set as below for all participants:

-        A single numeric calculation and answer section task

-        An audio-based picture task, with a one spoken sentence presented

-        A ranking task, with a set of five statements presented (based on the audio sentence prior)

Each participant was exposed to stimulus from the three Arabic varieties: MSA, EA & GA under the same conditions. All the verbal-guise exposed stimuli were audio and were followed by two tasks: a picture selection response task, then an attitude task. Additionally, a cognitive load was conducted within each trial, to maintain a status of cognitive stress, by using up the WM (see *2.2.3 The Contextual Factor of Stress*), whilst also performing as a distractor task. The distractor task was presented first to maintain equal cognitive load for the presentation of the audio stimulus, as the cognitive load was always placed prior to the audio being playing.
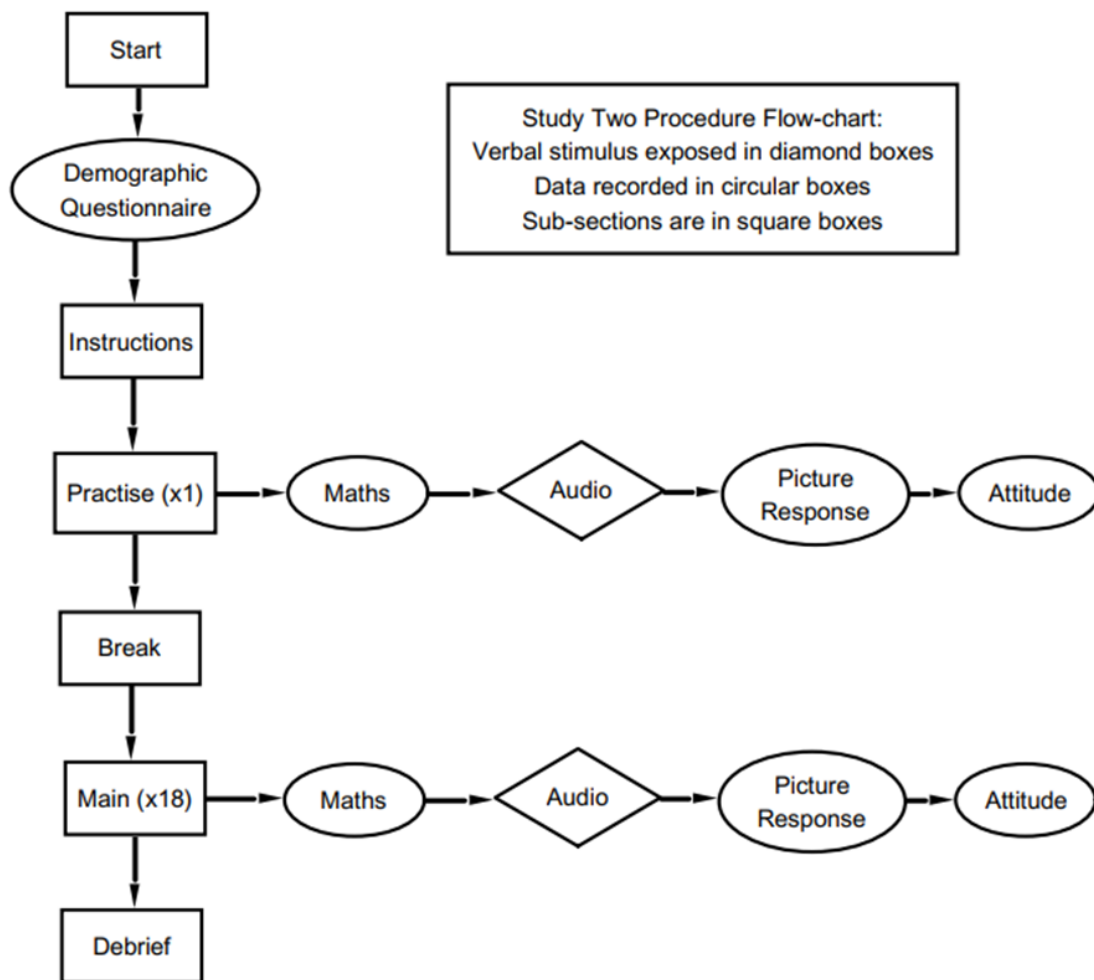
Figure 15: Experiment procedure flow-chart

5.6.1 Number Calculation Task:

The participants were presented a mathematical question for 3000ms. After the time

elapsed, another screen was presented with a selection of 4 possible answers for the

question, in a crosshair design (see figure 16). One was correct & the other three

incorrect. The placement of the correct answer was randomly generated by Gorilla.

The participant was asked to select the correct answer to the equation previously

seen. Furthermore, the order of stimulus and difficulty was also randomly selected by

Gorilla's randomiser; this was deemed acceptable given that the task functioned to

create cognitive load as a constant state, and the consistent placement of the task was the primary consideration taken to ensuring the state of stress was continual; as such the potential impact of ordering effects was considered subsidiary, particularly as the study's focus was on the intelligibility and trust tasks.

Figure 16: Presentation of maths response options

5.6.2 Picture Response Task

Participants were presented with the birds-eye map (see Visual Stimulus) whilst listening to a sentence (see Auditory Stimulus). Upon completion of the sentence, a screen appeared with four images of place markers from the map. These images were presented in a crosshair style. The participant was asked to select the place marker which accurately matched the directions in the auditory stimulus. The three incorrect images were based of place markers that would be reached, if one part of the instructions was incorrectly followed (see figure 17). There was a response window of 5000ms, after which the trial moved on, with an incorrect score recorded. The reaction time and accuracy of response (correct/incorrect) was recorded by Gorilla.
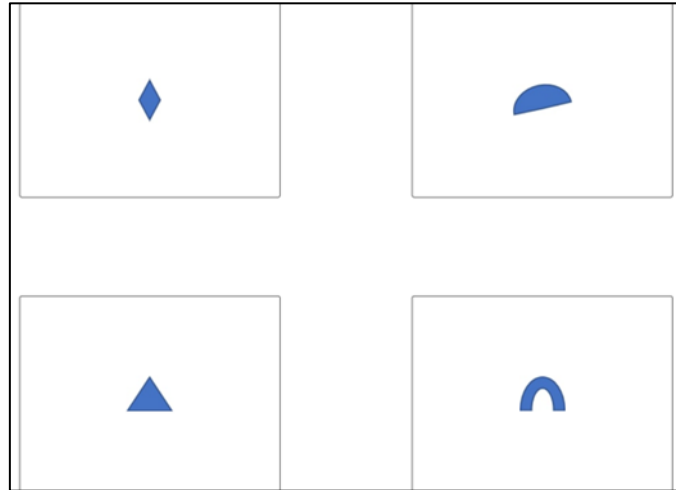
Figure 17: Presentation of buildings for response task

Auditory stimuli were arranged in a Latin-square-like design, whereby each dialect was presented in a fixed order, but the sentences within the dialects were randomised. The 18 sentences presented were equally divided across the three Arabic varieties; so that in total 6 sentences were presented from each variant (6 Egyptian, 6 Gulf and 6 MSA). There were three speakers used per Arabic variety, and two sentences from each speaker were used in the experiment (3 speakers x 2 sentences x 3 Arabic varieties) (see appendix D for stimulus used).

The presentation order of the Arabic varieties was also pre-set, using three staggered lists. This approach replicated a Latin-Square design, which counteracts any potential ordering effect by presenting the stimulus groups in alternating order equally across the participant groups (see table 32). Between the participants group, an equal number will be presented the stimulus in the order of reach list (20 participants x 3 presentation lists). This study uses a 3x3 design, as there were three different stimulus groups (Arabic varieties), and so the Latin square was adapted to a three-list ordered design.

Table 32: Presentation order of Arabic variants for picture response stimulus task.

	List 1	List 2	List 3
Repeated pattern	MSA	GULF	EGY
	GULF	EGY	MSA
	EGY	MSA	GULF

5.6.3 Attitude Task

The participants were presented with 5 statements accompanied by sliding scales.

These statements were presented across two screens, with 3 statements on the first

screen and 2 statements on the second. Each statement was relating to the

sociolinguistic element of solidarity. The statements, in order of presentation were

(the social factor is emboldened):

1. This speaker is **trustworthy**

2. The speaker is **reliable**

3. The speaker spoke **clearly**

4. The speaker is **honest**

5. The speaker is **respectable**

The participants were asked to rate their agreement with the statements by selecting

a point on a sliding scale between two polar points of Agree & Disagree. The sliding

scale was 0-100 and allowed for any point between the two endpoints to be selected,

such as 73, rather than just 70 or 80. Statements 1,3 & 5 were presented with Agree

on the left hand-side, with statements 2 & 4 being opposite with agree on the right

hand-side. This was to counterbalance any bias towards one side of the screen. The

points on the sliding scales were recorded by Gorilla.

Similar to the picture response task, the presentation order of the sentences was pre-

set. Using a Latin-square design, this time in a 5x5 form (see table 33). This

produced 5 presentation lists, which set the order of the trust statements and scales

for each question from that list. The participants were shown different lists, with 1/5

being shown each list; this was to disperse the lists across the participants in a

uniform pattern. As such, there were twelve participants for each list (5 lists x 12

participants = 60 total participants).

Table 33: Presentation order of Arabic varieties for trust questions

	List 1	List 2	List 3	List 4	List 5
	1	2	3	4	5
Repeated	2	3	4	5	1
Pattern (per	3	4	5	1	2
sentence)	4	5	1	2	3
	5	1	2	3	4

5.6.4 Presentation Method – Electronic

The tasks were run using Gorilla Experiment Builder (2020), an online experimental

study service provider (Anwyl-Irnive et al, 2020), allowing the task to be completed

anywhere, via phones, computers, or tablets. For this experiment, the restriction of

platform usage was set as only for touch-screen display devices, such as tablets and

phones. The fixed-response method was set to avoid any compounding effect of completion; for instance, (Pronk et al., 2020) identified that there are delay difference between the method of response, with touch screen providing a lesser delay than a PC and mouse.

There were two sections to the experimental stage procedure: a practise and a main. The practise contained a single trial; whereas the main contained 18 trials; with 6 trials containing verbal stimulus from each Arabic variety (3 Arabic varieties x 6 sentences). There was a break between the practise and main, to allow the participant to fix any technological issues & prepare for the main session (see previously presented figure 15 for visualisation).

On average, the study took less than 15 minutes to complete. A full debrief of the aims of the study were detailed upon the completion of the final trial; also, to withdraw the participant could exit the browser tab used for the experiment. There were no withdrawals from this study.

There were three tasks overall per trial, also once each task was completed the next began immediately; thus, there was no break in the main session. The order of tasks was pre-set and stimulus order for the picture selection task and the trust task were also pre-set; whereas the stimulus presented in the maths task was randomised by Gorilla.

5.7 Results

Overall, there were 1080 trials across 60 participants. Within which contained 360 trials for each of the 3 Arabic's. There were 360 reaction times and accuracy scores

for each Arabic in the picture response task. Furthermore, in the attitude task there were 5400 ratings across 5 social factors with in each of the Arabic's there were 1800 ratings. Additionally, there were 1080 reaction times & accuracy scores for the cognitive load task.

5.7.1 Data (pre-test) processing

The first stage in sorting the data was to calculate the average RTs and ACC percentage for each participant for each Arabic variety, additionally, the total trust score was also calculated.

5.7.1.1 Accuracy Rate

The accuracy rate was calculated using the average mean multiplied by 100, this was due to the accuracy being a binary scoring (categoric). This produced an accuracy as a percentage, rather than a score between 0-1:

$$Per\ particpant\ \left\{ \frac{Correct\ Answers}{Number\ of\ sentences} \times 100\ = Accuracy\ Rate\ (\%) \right\}$$

5.7.1.2 Reaction Times

A single dataset of RT was collated for analysis. This data set contained the simple mean for all RT's recorded in the study per participant when the response given was accurate. In the dataset, the times were sorted according to which Arabic variety was exposed in the trial, thus the dataset contained the scores split into three groups (Gulf, Egyptian and MSA).

5.7.1.3 Participant RT average

This dataset was produced by collating all of the RT recorded in total, which was 1080, (18 sentences x 60 participants), then removing all RT values where, in the trial, the response was inaccurate (scored as 0). A total of 695 values (233 Egyptian, 232 Gulf and 230 MSA) were retained in this reduction, which were then grouped by which Arabic variety was exposed. Which further reduced the dataset into 180 RT scores overall (1 reaction time x 60 participants x 3 Arabic varieties There was a further reduction by calculating the average RT of all accurate scores per participant for each Arabic variety:

$$Per\ participant \left\{ \frac{\Sigma(RT\ when\ Acc = 1)}{Number\ of\ RT\ when\ Acc = 1} = Average\ Accurate\ RT \right\}$$

5.7.1.4 Combining Trust

The reliability of trust questionnaire was calculated using Cronbach's alpha co-efficient; a test which assesses whether the responses to each question were consistent overall; to provide a value for the reliability score of the method itself (Cronbach, 1951). The values for Cronbach range from 0-1, with higher values indicating greater reliability, although this decision of acceptable reliability is scaled. Both Kline (1999) and Cortina (1993) categorised scores below .8 as unacceptable; with scores between .8 and .95 as good. Values of 1, or close to 1, infer minor variation, which is *too* perfect a score, and thus is deemed unacceptable (Kline 1999 & Taber, 2018). Konting et al. (2009) defined scores above 0.9, as excellent indicators of reliability, however, Taber (2018) warned that scores closer to .95 than .9 could infer unacceptability high reliability. Therefore, the acceptable alpha value

range for this thesis, is between .8 and .925. The reliability analysis was conducted, using Cronbach as the statistical model producing an alpha co-efficient value of .908, thus the reliability of the trust scores was acceptable (nigh excellent), allowing for the combining of the questionnaire scores into a single value as variation between measure was acceptable. This high reliability justifies the counter-balancing efforts taken to reduce the listing effect, such as variating the presentation order of questions and reversing the semantic scale at regular intervals.

There were 6 Arabic sentences heard per variety, each of which was followed by a set of 5 trust questions, so each participant provided 30 trust scores for each Arabic variety (6 sentences x 5 questions). In total, there were 90 scores recorded for each participant (30 scores x 3 Arabic varieties). The scoring scale of the questions ranged from a minimum score of 0 to a maximum score of 100, and when combined the total maximum score per Arabic variety was 3000 (6 Sentences x 5 Questions x max score of 100). This calculation sorted the overall trust dataset from 5400 data entries (30 trust scores x 60 participants x 3 Arabic varieties) to 180 (1 average trust score x 60 participants x 3 Arabic varieties).

Overall, post-sorting produced 3 datasets: 1 for the Accuracy rate, 1 for the accurate Reaction Times and 1 for the Trust scores (see table 34). In total, there were 540 values across the three datasets (180 values x 3 datasets). Each dataset contained 1 value per variable per Arabic variety totalling 180 values in each set (1 value per variable x 60 participants x 3 Arabic varieties). For consistency, each value was labelled by the Participant Identification Number, to ensure that each dataset contained the result of each participant in a way that could be identified in the event

of an outlier (of which there were none identified), or post-testing request for data removal (of which there were none).

Table 34: Allocation of data per variable for each Arabic Variety exposed

	Data per variable		
Arabic variety	Accuracy Rate	Trust	RTs
Gulf	60	60	60
Egyptian	60	60	60
Modern Standard	60	60	60

5.7.1.5 Testing Groups

Given that the demographic data (age & sex) was categorised into two groups per variable, an analysis was conducted to assess whether the data could be considered as a single dataset, or whether separate groups had to be considered on the basis of demographic factors. A MANOVA was performed on the Participant RT, ACC, and Trust scores, with Age and Sex positioned as fixed factors; this was to assess whether either factor had an effect on the dependent variables. There was no effect identified for Age ($p = .691$, $F = .489$, $n2 = .008$) or Sex ($p = .798$, $F = .338$, $n2 = .006$).

5.7.2 Descriptive Statistics

Following the data checking, based off the demographic information, the means and standard deviations for the three of the datasets were calculated Accuracy rate, Trust and (Accurate-only) RT for each of the Arabic's (See table 35).

Table 35: Descriptive statistics (mean and standard deviation) by Arabic variety

Arabic variety	Accurate Reaction Time		Accuracy		Trust Score	
	Mean (ms)	Std. Deviation (ms)	Mean (%)	Std Deviation (%)	Mean (%)	Std Deviation (/3000)
Gulf	5366.92	1202.46	65.03	19.04	1786	119.5
Egyptian	6004.86	1516.10	58.70	15.02	1968.28	164.51
Modern Standard	5639.67	692.97	63.94	21.90	1754.7	124.61

5.7.3 Inferential Analysis

This analysis will report on the variables in the following order: Accuracy rate, followed by RTs of accurate responses then Trust scores.

To investigate the relationship between the accuracy rates, RTs and the trust scores between the three Arabic varieties, a repeated measures MANOVA was performed on the dataset, finding a main effect for RT ($p = .014$, $F = 4.365$, $n2 = .047$) and Trust ($p < .001$, $F = 42.109$, $n2 = .322$). There was no effect for the accuracy rate ($p = .173$, $F = 1.773$, $n2 = .020$) (see figure 18).

A post-hoc Bonferroni test was conducted to examine the relationship between the three Arabic varieties, with both the RTs and the Trust scores, to identify where the differences occurred (see figure 19 for RT, and figure 20 for Trust). Firstly, the RT will be reported, followed by Trust.

For the RTs, there was a significant difference between Gulf and Egyptian ($p = .011$, SD = 216.66). No significance was identified in the other two comparisons: Gulf Vs MSA ($p=.629$, SD = 216.66) and Egyptian Vs MSA ($p = .281$, SD = 216.66).

For Trust, there was a significant difference for two of the relationships: Gulf vs Egyptian ($p < .001$, SD 25.14) and Egyptian Vs MSA ($p < .001$, SD = 25.14). There was no significant difference between Gulf and MSA ($p = .664$, SD = 25.14).
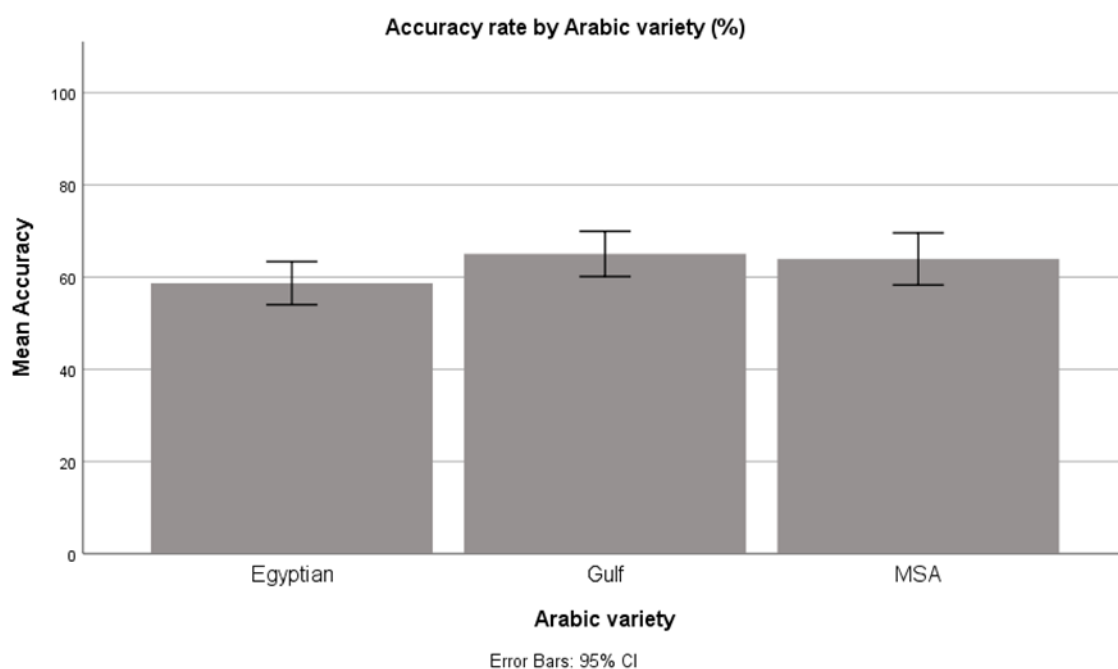

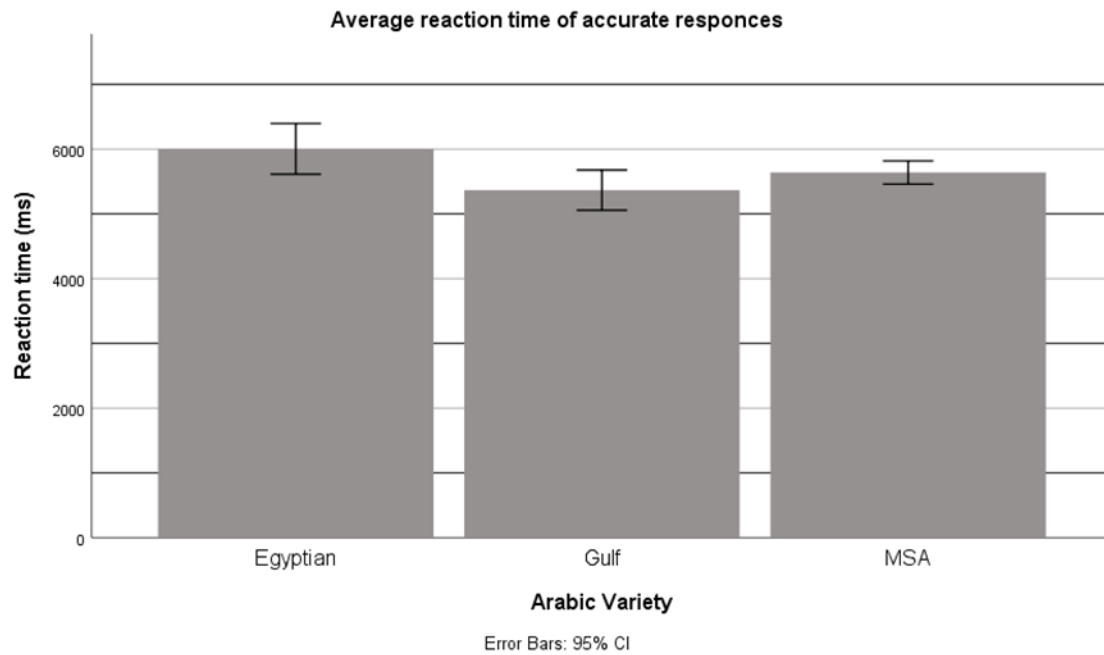**Figure 18**: Accuracy rate per Arabic variety

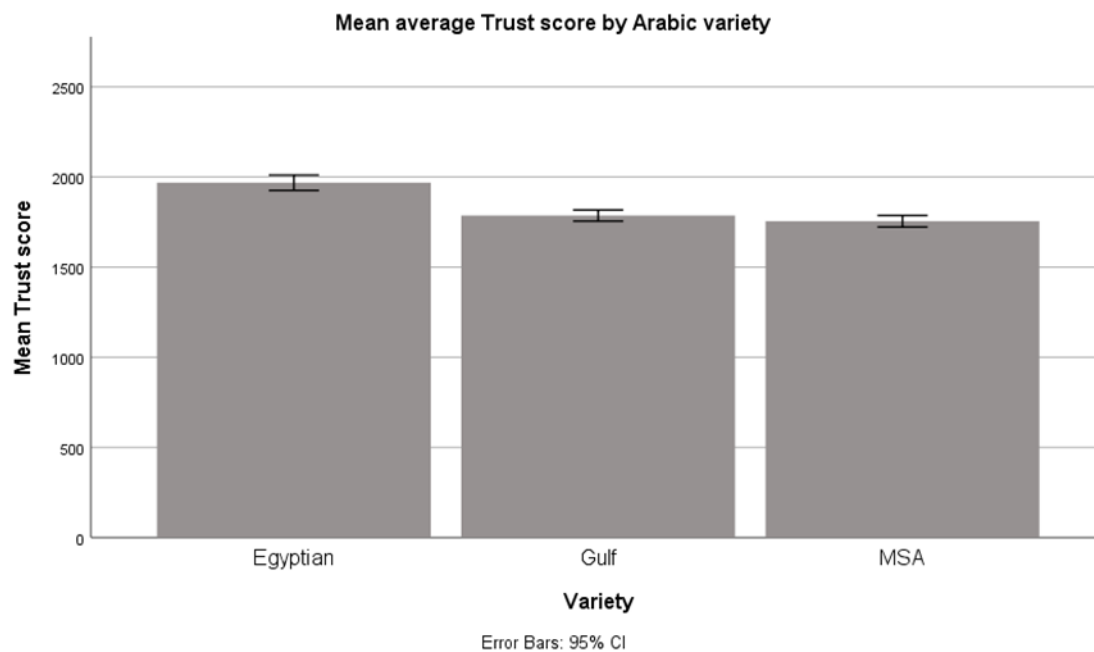**Figure 19**: Average reaction times of accurate responses by Arabic variety

Figure 20: Average Trust scores per Arabic variety

5.8 Discussion

The results of this study address the three hypotheses presented as well as

informing on the overarching issue of language attitude measuring to indicate

underlying social trust, and the extent to which the diglossic status can aid in the

predicting of trust levels. These results build upon the previous study, with the

production of additional evidence on the intelligibility between the three Arabic

varieties on Saudi Arabians. In a similar fashion to the previous studies, the structure

of this section will be as such: discussion of each hypothesis formulated, in order of

presentation; followed by a discussion on the two overarching issues: language

attitudes & social trust, and diglossia in Arabic. The latter issue will be considered in

relation to the context found in TWB's (2017a) report, where individuals did not trust

information provided by members of either the same nation, or of neighbouring rival

nations, favouring written MSA instead. This discussion will explore whether the trust

issues are related to the prestige of H-varieties in contrast to the negative traits

associated from the L-varieties, in which case the diglossic status is a defining

feature for trust. To recap, the three hypotheses were:

1. The speakers using the localised QA will be viewed as the most trustworthy,

   as the associated attributes support the identification of the speaker as part of

   the In-group socially

2. The trusting responses towards MSA speakers will be more positively scored

   than towards non-localised QA's, as the established social grouping supports

   viewing MSA speakers as within the community, and other QA's as outside

3. The viewed trustworthiness results between Arabic varieties, in terms of

   performance, will mirror the intelligibility relationship, i.e., the most intelligible

   will be the most trusted and vice versa, particularly for low intelligibility, for if

pp. 383

communication cannot be understood, the veracity of the message equally

cannot be assessed, therefore, trust will be low

5.8.1 Hypothesis 1

The results of this study reject this hypothesis, as the EA trust scores reported were

higher than the GA scores. In fact, the results of GA trust were statistically

homogenous with MSA scores, with lower responses when compared to the EA

results. The rationale for this hypothesis was rooted in the familiarity effect, when the

relationship of trust is directly related to the exposure of a language variety;

particularly when the discourse partners are part of the same social In-group. As

such, it was proposed that the GA speaker would be perceived as the most

trustworthy as the listeners were from the speaking population; however, the results

do not support the assumption, that social In-groups are a primary factor in trust

when Arabic speakers are considered. From the literature, two explanations for this

were identified. Firstly, that there is deeper than expected distrust within Saudi

society towards members of the same nationality, similar to the Saudi vs Iranian

relationships (see Chakrani (2015)) for an assessment of the influence of negative

language attitudes on language accommodation and trust). Historically, the Arabic

speaking nations have existed in a sense of rivalry, in a geo-political sense, resulting

in negative stereotyping in both nations (Versteegh, 2014); an effect which may have

permeated into the general society in Saudi Arabia. The second explanation may be

due to un-noticed Western bias from previous studies, as the primary support for the

familiarity effect identified in the literature review focused on English usage in

America (Clopper & Bradlow, 2008; McCloy et al., 2015), and there was an

assumption that the perception of social groups based on familiarities was universal across language groups and speech communities. In fact, the rejection of GA supports Fiedler et al's (2019) findings, which are that the social expectations and reception of intelligibility differences and language variation depends on the nationalities involved, both of the speaker and of the listener. Furthermore, the judgements for mixed-nationality or mixed language variety situations are based off several factors; including social cohesion, familiarity and socio-political divisions being identified as part of the mixture that forms underlying biases. Overall, the results support further research into the social factors which, in part, form the attitudes that are attached to the language use of speakers both within the social groups, and between different groups.

5.8.2 Hypothesis 2

From the results this hypothesis is rejected, as MSA was the least trusted variety, and the non-localised EA was scored as the most trustworthy. The lack of perceiving EA as untrustworthy counters the expected result from the literature, and indicates that the prestige associated to MSA, and GA to some extent, did not result in positive views of the speakers overall. However, the results could be indicating that the associated attributes were not singularly based on the national social group, rather the attributes were based on the organisational reputation. This explanation uses Coombs's (2007) Situational Crisis Communication Theory (SCCT), a model which predicts that an individual's behavioural response to disaster messaging is based on three core elements: the emotional response, the professional standing of a language community or organisation, and the prior relationship between the

stakeholders. The SCCT identifies that negative prior relationships and attributes will counter-act the effectivity of disaster responses, as people will hold onto the negative views formed prior to disaster, and thus are more likely to respond in opposition to the orders given. If SCCT is applied to the Arabic varieties used, the trustworthiness of EA may solely be due to a lack of negative prior relationship between the Egyptian communities and Saudi's; or if there is a negative prior relationship it is minor in comparison to the disputes or differences recorded for the average MSA and GA user. The presence of underlying negative views between Arab populations is not uncommon, for instance Mirshahidi (2017) explored the relationship of attractiveness and honesty harboured within linguistic attitudes; a study which identified that ethnic groups across the Arabian peninsula's often view neighbouring ethnic groups as rivals to be distrusted.

5.8.3 Hypothesis 3

This study's results support the rejection of this hypothesis, as the relationships between the trust scores did not correlate to the intelligibility scores. There were two intelligibility scores, both of which possessed different relationships between the Arabic varieties than the trust scores (see figure 21 for a visualisation of the relationships).

For clarification, as previously established, EA was found as the most trusted and GA and MSA equally the least (due to homogeneity); yet this dynamic was not equal for accuracy rates or reaction times. With accuracy rates, there were no differences detected between the responses to all Arabic varieties, therefore, the expectation from this hypothesis cannot be realised as the mirrored relationship between

pp. 386

accuracy and trust would be when there was no difference in trust levels. Similarly, with reaction times, the relationship was that GA was the fastest responded to, and EA the least; with MSA being somewhere between GA and EA. If the mirroring between trust and reaction times was realised, it would require EA to be both the fastest and heterogenous to MSA responses, and GA and MSA would need to be slower than EA, but equal to each other. Additionally, an inverse mirror relationship was also not present. For the inverse mirror to be true three relationships would be needed: EA to be the least intelligible yet the most trusted, GA the most intelligible and least trusted and MSA somewhere between the scores of both. However, the position of MSA in the relationships was not as clear-cut; as it was equal to both GA and EA (despite both being different from one another), indicating a middle-ground relationship for reaction times; however, with trust scores, there were differences identified between MSA and EA but not with GA. Therefore, the lack of mirror relationships indicates support for the notion that intelligibility is not directly related to the perception of a speaker as being truthful.
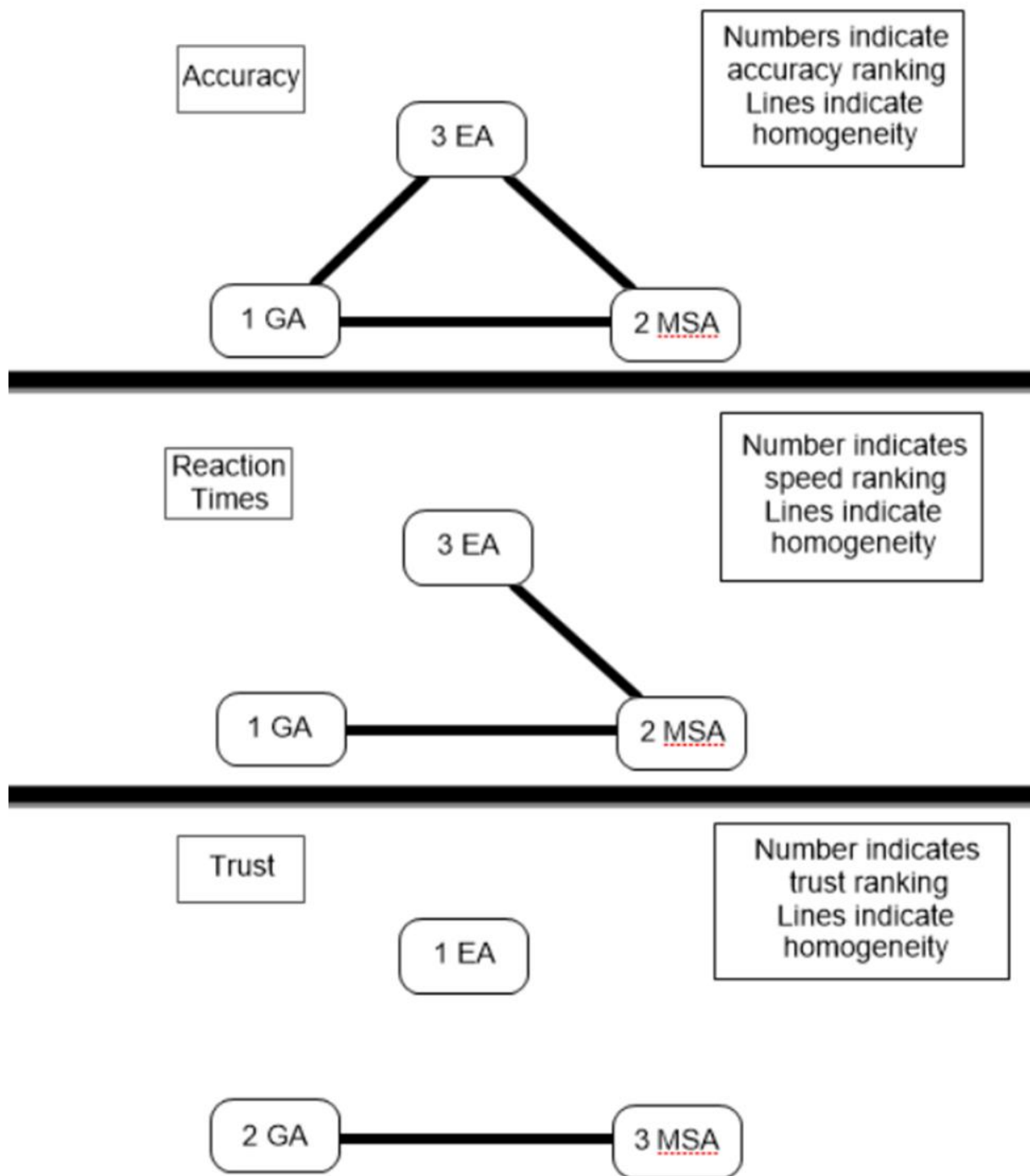
Figure 21: Relationships between Arabic varieties by accuracy rates, reaction times and trust scores

5.8.4 Overarching issue: Diglossic status

In this study, the relationship between the diglossic language varieties found in Saudi Arabia and the underlying trustworthiness associated with a speaker was explored. The results provided evidence against multiple sociolinguistic arguments, particularly the relationships between diglossic language varieties and the perception of speech in said language varieties (Djennane, 2014; Abdel-Rahman, 2016). The first argument disputed was that the prestige of a diglossic language variety will be attributed to a speech community (Walters, 1996; Albirini, 2011), in that the H speakers will be viewed differently (and more positively) than the L speakers. In the results, this study identified that both the H and the L language varieties were viewed similarly by listeners, thus contrasting to the expectation that the H speaker will be afforded greater social trustworthiness. In Arabic sociolinguistics, it is established that MSA is the language variety for use in the official contexts, which often require truthful discourses and an expectation of honesty (Høigilt & Mejdell, 2017), whereas Najdi, as a Gulf QA does not experience the same pressures (Khalil, 2012; Al-Rojaie, 2013) .

Despite the social expectations for H speakers in official contexts, the virtues of trustworthiness were not attributed or associated with the language variety usage. In previous attitude studies, the connection between social and linguistic attitudes has, by-and-large, indicated a mirroring relationship, whereby the higher the social trust, the more favourably the speakers are perceived (Albirini, 2011; Omar & Ilyas, 2018). If this trend is applied to the results of this study, it can be inferred that social distrust of the officialdom is higher than presented in the literature (see *2.3.5.3 Status quo of*

Ausbaucentrism and Diglossia in Mashriqi Arabic social attitudes). There is evidence of social distrust within the neighbouring regions to Saudi Arabia, such as the Israel-Palestine conflict and the Yemen civil war, which are two conflicts that the Saudi Arabia is actively involved with, whilst denying responsibility for any results (Issaev et al., 2022). This may have impacted the trust attributed to the state by the local population, so much so, that the veracity of the state (which is an absolute monarchy (Bowden, 2019)) is placed into question. Across the Arabic speaking nations in Northern Africa and in the Middle East, there has been widespread disorder since the beginning of the Arab Spring, and for several nations there was governmental change; however, for Saudi Arabia, there has been less change when compared to neighbouring states (Bowden, 2019). As such, the underlying attitudes in the population may also still harbour negative attitudes towards the government due to a lack of change (ibid). Ultimately, the two reasons for the rejection of trust towards the H variety are indicative of wider social change, or the lack thereof.

The impact of social change on the diglossic stability in Saudi Arabia is uncertain, however, there is evidence that modern Saudi's when using social media, will use their local QA (Sadat et al., 2014), such as Nijari for posting content and responding to current affairs (Abdul-Mageed et al., 2018). This expansion in use indicates that the L diglossic varieties are acceptable for use in the wider social context, whereas the use of the H variety is far lesser. Given the prevalence of social media in Saudi Arabia, the stability of the diglossic continuum can be pulled into question. If the H variety is losing status, in terms of contexts of use (Abdel-Rahman, 2016) or prestigious attitudes (as indicated in this study), it can reflect the rejection of the H as

the defacto language variety of use in the wider environment. In itself social change in Saudi Arabia may be impacting the diglossic status of Arabic, which is a case touted as the best example of stable diglossia (Snow, 2013), therefore suggesting that diglossia is only stable when the L does not expand, which is a view widely supported in sociolinguistics (Kaye, 2001, Saiegh-Haddad & Lina, 2018).

An additional aspect in the Saudi diglossic situation is the media-dominant EA, which acts in a pseudo-H variety position (Samin, 2012; Al-Suwaiyan, 2018). The delivery of EA may be through media, but within the narratives depicted, the context of use for EA involved many which are exclusive to MSA in Saudi Arabia, such as in court cases and with politian's speech (Kraidy & Khalil, 2018). This controlled image of EA usage may have influences that associated characteristics with the speech community. This study's results indicated that EA speakers were noticeably perceived as more trustworthy than GA and MSA. A rationale for this result can be found when considering the stereotypes developed towards the three varieties in Saudi Arabia. As noted above, the attempts at social change, particularly in terms of government style, largely failed; however, the dismissal of protests was violent, and the police force and officials of state were responsible for the dispersal of unrest (Bowden, 2019), as such the stereotype of officials may be that they are corrupt or against the people.

Furthermore, Najdi Arabic speakers are associated with the general population, and the stereotypes for the speakers are not specifically linked to the GA speech community in this case. Remember that this study assessed language attitudes of Saudi's, so the negative stereotypes towards GA speakers would have indicated how

a listener perceives their own community, as the participants were GA speakers themselves. As such, the reputation and stereotyping of fellow GA speakers based on using GA is unlikely to be negative, due to a shared In-Group (Tajfel, 2004; Chakrani, 2015). In reality, this social relationship indicates negative stereotyping towards the MSA users, and neutral towards the GA, but given the limited exposure of EA, a positive stereotype is more applicable.

In emergencies, as Coombs (2007) SCCT proposed, the prior relationship between the associated groups can impact the trustworthiness of messaging. With this, the lack of explicit negative stereotyping towards EA speech may be an indication as to why EA was the most trusted; as the relationship is reflective of a lack of negative stereotyping towards a speech community rather than the diglossic prestige of the language variety.

The context of this study (disaster evacuation orders and stress condition) can provide further insight into the relationship between diglossic status and trust, noticeably, the familiarity effect on developing urgency in a message (Sandman 1989; 1993). Within Saudi Arabia, there are two primary language varieties used, the H of MSA and the L of Gulf QA's, and the exposure of EA is via passive engagement (Bassiouney, 2009; Frishkopf, 2010; Samin 2012). It is unremarkable when GA or MSA are used in contexts in the wider Saudi community, as both hold regular function in discourses; the contexts of use may vary, but the use of the language varieties themselves is commonplace (Al Suwaiyan, 2018). As such, when hearing speech of either Arabics' is not a cause for attention beyond the normal effect given. However, EA is uncommon in usage in the nation, and it is generally restricted to

programs from traditional media, such as TV shows (Frishkopf, 2010). In emergency

contexts, an important aspect needed for compliance with evacuation orders is

urgency (Sandman, 2012), for if the population believe that the risk requires a rapid

response, they are more compliant with messaging and evacuations. It may be that

the hearing of EA, with the emergency contexts, was registered as ab-normal, and

therefore indicative that the situation was out of the ordinary. As it is important to

remember that the speech community of EA is perceived as an Out-group to the

general Saudi society, with the use of MSA and GA actions indicating the speaker is

part of the In-group. Therefore, the use of an outsider Arabic could explain why EA

was perceived as the most trustworthy due to the perception that EA use is an

abnormality. Unexpected or uncommon language use can increase the perceived

urgency and severity behind the message, particularly when the standard

conventions with the speech community are ignored (Sandman, 1989; 1993), as

listeners pay more attention to the details in the messaging, particularly in

emergencies. An increased attention provides greater social weight to the details of

the messaging (Sandman, 2012), and in this study, the details were evacuations and

the map presented alongside, which indicated a risk though a visual danger sign.

This combination of language use and in a replication an emergency context,

provides support for the fieldwork findings, in that the greater the attention obtained

for a message, the stronger the urgency perceived. Therefore, supporting the use of

an Out-group language variety, if the goal is to obtain attention and improved disaster

responses.

5.8.5 Responding to research question 3

(3) Is underlying distrust between Arabic speaking communities a factor to account for in disaster communication planning?

When considering whether language attitudes towards a speaking community have impact on the perception and responses to messengers from each community in question, the first element explored was whether the message itself was received accurately. In this study, there was little difference in terms of understanding between all three of the Arabic varieties tested; indicating the results of the study are not impacted by the accuracy of understanding of the message, therefore any distrust is based predominantly on the characteristics and stereotypes that could be attributed to a speaker. The results for the trust questionnaire indicated that there is a noticeable difference in terms of trusting an Arabic speaker, particularly when the speakers are from a different social group. The speakers of the local Arabic's, in relation to the listeners, were perceived as significantly less trustworthy than the speakers of the non-local Arabic. There was little difference identified in the trusting towards the local QA when spoken as well as to the local standard Arabic (gulf accented MSA), indicating that if the communication is distributed within the same society, and social group, then there is little evidence to support accounting for trust differences within the population based on which Arabic is used. However, if an international speaker of Arabic is used, then they will be perceived as being more trustworthy than the local speakers, therefore, the attitudes are a factor to consider for a disaster policy.

The social prestige was not a primary factor identified for indicating the trustworthiness of a speaker based on which Arabic variety they use. The literature supported the considering of H varieties of a diglossic situation as being more trustworthy (Walters, 1996; Chakrani, 2015), due to being associated with contexts where high veracity and honesty are required for the conducting of business, such as in the judiciary and within government. Furthermore, the use of L varieties in the domestic contexts was not identified as a factor to impact the trust, provided the speaker was identified as being within the same social community. This study found that the further away a QA is from the listener, in terms of social status and context, the greater the trust, a result in contrast to the literature.

Moreover, the rationale for why the outsider's QA was more trusted, may not be a result of the QA itself, or the social status of the QA (which is L in the diglossic continuum); in fact, the familiarity with the language used may impact greater. Coombs (2007) highlighted that the use of unfamiliar language conventions and varieties in an emergency improves the response to the delivered messaging, as the hearing of the outsider language variety is treated with greater attention by listeners. This attention increase is rooted in the risk-perception of the population, as the unusual language can be attributed to extra-ordinary contexts or developing situations, which in turn increases the perceived severity. This causal sequence may be more important for the development of trustworthy messaging in the Arabic speaking continuum.

The research question related to whether trust is a factor to account for, to which this thesis takes the position that it is. However, this may not necessarily due to the

diglossic status of the language varieties involved, but more so to do with the reception of critical messaging in a disaster context.

5.9 Conclusion

This chapter explored the underlying language attitudes of Saudi Arabians in relation to the trustworthiness of Arabic speakers based on which variety of Arabic was used. The findings highlighted that the status quo of attitudes and unity within the Arabic speaking community is not uniform across the Arabic's used in the experimentation. The QA of the outsider community (the Egyptians) was identified as being the most trusted, supporting the familiarity (or lack thereof) effect (Gass & Varonis, 1984) as a primary factor in improving the attention given and urgency developed from an emergency message. Furthermore, this study identified a heterogenous relationship between the local QA and MSA within Saudi Arabia, therefore refuting the notion that the diglossic status of an Arabic variety is indicative of the social trust of a speaker (Walters, 1996). Despite MSA being the H variety, and associated with contexts where high veracity is essential, the speech community does not associate the speakers of the H variety with such values any more than they view the local L variety. The application of this result infers that the political ideologies held by Arabic scholars, particularly that Standard Arabics are indicators of high social prestige and trust (Bassiouney, 2009, Versteegh, 2014 & Abdel-Rahman, 2016), are less representative in the modern era.

Chapter 6: Discussion, limitations, and conclusions

6.1 Summary of findings

This thesis contained three studies which explored three separate linguistic aspects of language planning and policy that are specific requirements for developing effective disaster language policies and communication strategies. Each aspect was explored with a guiding research question, separated across chapters 3 through 5. Chapter 3 investigated the usability and acceptability of using pre-existing international language policies for disaster contexts, in this case the UNLP, finding that the policy is better suited as a guiding example of standard for language resources and stockpiling rather than as a default policy to deploy in a disaster. A new method (Reach) for quantifying whether a language is acceptable was presented and used to identify the ideal languages to use in a disaster situation, which was compared against the UNLP for impact potential. Finding that the UNLP was overall unsuccessful as a single policy for facilitating language support and communication if used, however, the results indicated that expanding on the UNLP with localised majority languages could provide marked improvements to language access. Within this chapter, the issue of Ausbaucentrism (Tamburelli, 2014), and the severity of such bias, that exists in both international and national level languages policies, was addressed, with results highlighting that Ausbau-led language policies are incompatible for deployment in a disaster. As Ausbau-defined languages do not account for the communication between speakers as a vital element, the selection of languages in a disaster policy must be careful of assuming that speakers of an

established language can understand one-another. The extent of Ausbaucentrism in the global language demarcation system was identified as an attributive factor that explains the poor status of communication reported in TWB (2017a).

Chapter 4 provided a study which presented a method quantifying the intelligibility between three language varieties which are classified as being part of the same Ausbau language but have also been recorded as being separate languages by Abstand methods. The findings supported the classification of Najdi Arabic and Cairene Arabic as separate languages from each other but rejected the classification of MSA as a stand-alone language too. The results indicated that the ideology that MSA is a go-between Arabic variety for inter-national discourses (Albrini, 2011) may hold merit in the terms of communication and Abstand methods. The rational for the use of MSA presented in the literature was predominately from an Ausbau perspective, therefore, this study highlights how Ausbau and Abstand classification do not always deviate, and it can be reflected equally in some cases. That being said, the findings also highlighted an issue within Abstand research, that of a disagreement as to the threshold for declaring a language separate from similar language varieties. This was evidence by the average accuracy of responses to sentences in varying Arabics ranging from a low of 58% to a high of 72%. The variation both between the Arabic varieties tested and between the two studies, which measured intelligibility of Arabic between Saudi Arabian listeners, indicates that the intelligibility relationship between the varieties is variable dependent on the language units heard. This thesis did not control for specific linguistic features for measuring intelligibility, as the focus was providing insight into the relationship

between Arabic varieties and communication within simulated disaster contexts to indicate the status quo of intelligibility in language use. Furthermore, the results of this thesis further support previous mutual intelligibility studies by finding that total mutuality in intelligibility is not a characteristic of natural language processing (Tang & van Heuven, 2007; Kürschner et al., 2008; Gooskens & Swarte, 2017; and Gooskens et al., 2020). This thesis supports the position that mutual intelligibility thresholds should not be arbitrarily set at total intelligibility (as in 100% accuracy) and supports Tosco & Tamburelli's (2021) position that intelligibility should be viewed as a naturally variating continuum rather than a fix-point continuum.

The final study, in chapter 5, provided insight into the status of social trust towards speakers based on the language variety spoken, by measuring the underlying level of trustworthiness of a speaker in a verbal guise design experiment. The findings dispute the assumption in sociolinguistics that the diglossic status of a language variety is indicative of the social trustworthiness of the community (Ivanova, 2013; Ready, 2018); as the trust scores were highest for the non-local QA of Cairene Arabic. In addition, the lack of trust towards the speakers of the local QA and SA indicate support for the familiarity effect (Sandman, 2012) and the impact of reputation on the perception of trust, validity, and authority (Coombs, 2007). Furthermore, the results can indicate that the status of MSA in Saudi Arabia, as the prestigious and official H variety is possibly weakening in default trustworthiness in the local population. Although, this may be due to the dominance of GA on social media (Abdul-Mageed et al. 2018) or due to the reliance of MSA on the local QA in

the nation which impacts the familiarity of processing SA (Saiegh-Haddad, 2003, 2004 & 2007 and Asadi & Abu-Rabia, 2021).

6.2 Implications and Impact

The initial goal of this thesis was to investigate the status quo identified in TWB (2017a), to provide insight into whether the situation was unique or a result of wider sociolinguistic issues. From the results of the three studies, this thesis takes the position that both the lack of communication and the language barriers are a result of the ongoing issues in sociolinguistics, in terms of research focus, underlying biases, and poorly justified assumptions. The first impact of this study is the evidence that the issues in the field are in part connected to language policy issues, bridging the gap between the research issues (which are particularly ignored in the Arabic language research field) and their impacts on the vulnerable populations displaced from conflict zones.

Furthermore, three methods were developed and presented in this thesis to measure the suitability of a language in a disaster policy. The first method (Reach) provides quantitative data analysis for justifying the selection of a language for national level disaster responses. The approach uses the density of language speakers within a target area and to justify which language to use based on the target population within the disaster region. This approach can be applied to both the nation state, but also to regional areas, providing that there is data on the speaker populations and languages within the target region. The advantage of the Reach approach is that it also requires the infrastructure for any selected language to be suitable for the

deployment in a disaster, therefore accounting for the deployability of a language within the method itself. Additionally, the Reach approach allows for objective comparison of similar disasters and language policies deployed in a like-for-like manner, as the required elements for comparison are simple enough to be universal across disasters and populations.

The second method presented allows for the measuring of intelligibility itself, which in this thesis was deployed alongside an environmental condition (cognitive load as a stressor). With the intelligibility test design, the overall understanding of a message could be assessed based on the accuracy of response and speed of reply. The general accuracy rate between speakers can be calculated, and therefore replicated for future studies as an established method for measuring intelligibility as an overall phenomenon. The test mixed the stimulus SPIN (Kalikow et al, 1977) test with the syntactic-variation design based on the TROG-2 (Bishop, 2003), producing an experimental design which can identify where the misunderstanding occurred within a message, as the image selection options vary based on differences in the perceived heard sentence. For clarification, in the example provided in chapter 4, one of the incorrect options was representative of the gender of the subject being incorrect 'he' when the correct option was 'she'. Further deployment of this method could be used to explore whether the placement of the variation impacts the overall intelligibility of a message in Arabic and other languages.

The third method utilised a verbal guise design alongside a contextually rich question context. The intelligibility of directions and underlying attitudes towards the trustworthiness of the speaker, allowing for inference of the perceived veracity of

both evacuation messages and the speaker who delivers the message. In disaster management, a key tenant is that messaging delivered is understandable (as in intelligible) and trusted, and this combination of a verbal guise mixed with an IELTS verbal reasoning direction test design allows for both trust and intelligibility to be tested in a single study. The results can infer directly to the situational responses in emergency communication and disaster management situations, as the stimulus is directly reflecting the demands and requirement of emergency messaging. Different messaging strategies can be tested within this design, to allow for empirical analysis of the varied structures, dictating how public announcements should be delivered, thus allowing for the identification of which aspects of language improve accurate communication.

Overall, methodologically, this thesis has presented an Abstand-led approached to language classification, and selection for language policies. Kloss's (1967) original proposition of Abstand highlighted the requirement for linguistic data-led approaches and techniques to classify language by Abstand, and this thesis adds to the ongoing efforts within sociolinguistics to operationalise Abstand as a standardised demarcation system of language. This thesis has provided a standardisable testing design for measuring intelligibility, which can be applied with and without set interspersed contextual factors. The first experimental design can allow for the pin-point assessment of where intelligibility is weakest, for instance, whether the unintelligibility is found based on the syntactic position of the information. This can be applied to further studies to assess whether the syntactic position of a word can impact the intelligibility, as measurable by using the incorrect answering to the target

word picture task. Stretching our understanding of intelligibility beyond the binary understand or not dichotomy provides additional depth to our understanding of the phenomenon itself whilst also providing research avenues for focused testing.

Additionally, the testing designs presented within this thesis can be used to measure intelligibility between a number of language varieties simultaneously, as this thesis compared three language varieties in each design, yet the test itself can be expanded to include additional language varieties. This application could be used in further research to explore the whole landscape of language varieties within an Ausbau language in a single design based on a spectrum of sample population from the language community groups. A streamlining that would also provide quantifiable evidence toward the rationalisation of the intelligibility thresholds to demark a language variety as a distinctive Abstand language itself.

Beyond the methodological implications, this thesis also provides insight into the classification of colloquial Arabic's and the extent to which Arabic, as an established international language is impacted by Ausbaucentrism (Tamburelli. 2014, 2021. This thesis supports the position that Arabic typology is currently unfit for use in linguistic applications and when considering language-based differences, due to the variation of Ausbau Arabic being an established issue that causes communication barriers. When examining Arabic, this thesis builds on the conclusions presented by Zbib et al. (2012) and Kwaik et al. (2018b), which is that of Arabic is itself a complex set of languages which are compounded together by a sense of socio-ethnic identity, yet this reliance of unity undermines efforts to develop our understanding of the language practises and linguistic differences of the communities across the Arab

regions. Furthermore, the results of this study highlight the complex diglossic continuum that exists in Saudi Arabia and challenges the ideology that the prestige of Standard Arabic as a H variety is indicative of the social trust a speaker will be associated with. This thesis presents evidence to reject this ideology and supports the role of frequency of positive exposure of a language variety, particularly, the extent to which it commonly appears in the wider community, as a greater indicator of trustworthiness than the diglossic social status. A view which deviates from Snow (2013) and Ivanova (2013), and therefore, this thesis adds to the debate regarding modern Arabic diglossia, however, this thesis does so from the under-researched context of disaster communication.

6.3 Limitations

Within this thesis, there are several limitations to note. In this section, the limitations will be separated into being either methodological or thematic and shall be discussed in presented order. The first limit of this thesis overall is that the Arabic's studied in chapter 4 & 5 are exclusively found natively in the Mashriqi region, and Arabic scholars have historically regarded the linguistics and socio-political differences between the cultures inhabiting the region as minor (see Cadora, (1989); Isaksson, (1996); Holes, (2005) for overviews). Whilst the results indicate evidence to the contrary, the scope of language varieties explored is limited to the colloquial Arabic's used in the capital cities of the geographic neighbours Egypt and Saudi Arabia. Furthermore, the participant groups for both chapter 4 & 5 were only Saudi's, and this coupled with the focus on Najdi Arabic, limits the inferences to suggestive of the Saudi sociolinguistic situation, and any change occurring within the communities of

Central and Northern Saudi Arabia. Whilst the results indicate a relationship within Arabic diglossia, the extent to which this context represents the wider Arabic diglossic status is limited by the slim volume of Arabic varieties chosen for experimentation.

Additional limits are related to the experimental chapters 4 & 5. The testing groups for both studies were small in terms of volume of participants, and whilst this was counter-balanced with increase trial testing of active participants, an ideal situation for the measuring of general intelligibility would be a between subjects' design, where there were more participants with less questions asked. Therefore, the results are indicative of the average response when processing information in bulk, rather than responding to a single message presented in solo, which could be representative of sudden emergency messaging more so than bulk trial testing as used in this thesis. Another limit is that the experiments conducted investigated the intelligibility and perceived trust of EA through a GA population, but did not test the reverse, I.e., GA through an EA population, therefore this thesis cannot provide insight into the overall intelligibility relationship between the two QA's. To recap Gooskens identified that the relationship between two language varieties is not reciprocal (see van Heuven et al., 2015; Gooskens & Swarte, 2017 & Gooskens & van Heuven, 2020), and thus this thesis cannot provide evidence as to the status of intelligibility relationship of EA and GA when EA is the population of listeners. The final methodological limitation is related to the application of cognitive load as a stress condition. Whilst CLT has been identified as a method to induce a heightened alertness and increase demand on the working memory, which results in a biophysical response and release of cortisol, as

outlined in *2.2.3 The Contextual Factor of Stress*. The extent of which stress was experienced by the participants was not tested within this study, and whilst the psychobiological properties of stress response were not the primary testing variable for the experimentations, the measuring of the stress response would have provided an indication of the confounding variable that stress could have on intelligibility. This is given that stress response tolerance and intelligibility are both measured on a continuum, and thus there may be an underlying relationship between the two factors, as stress reduces the working memory capacity and could impact the internal processing of linguistic differences between the Arabic varieties. This final limitation is not to say that the status of stress cannot be assumed in the studies, as previous literature supported that there would be a response (Klingner et al, 2010; Birkett, 2011; Bong et al, 2016), rather the limit indicates that testing for stress, using cortisol saliva monitoring could have provided a more detailed analysis of the relationship between stress experienced and the accuracy of similar language varieties.

With the thematic limitations, the first one identified is that only one local QA was tested (Najdi), a QA which arguably holds the highest prestige of the L varieties within Saudi Arabia, by virtue of the variety being the standard for discourse in the capital city Riyadh (Watson, 2007). Therefore, the closeness in social trust between Najdi and MSA may have been a result of the shared social power attributed to the capital city, which is a centre of socio-political power. Furthermore, the limitation is that this thesis cannot provide insight into the other diglossic QA's found in Saudi Arabia such as Hejazi or Khaliji. Another limit is the disagreement as to intelligibility rates reported between study's 4 & 5, with the former favouring the separation of GA

and EA based on a difference in accurate understanding (therefore lower

intelligibility), yet the latter providing support for the opposite, as there was no

difference quantified between all three Arabic's explored. Therefore, the application

of accuracy rates from this study should be limited to highlighting the complexity of

testing for intelligibility and the natural variability of intelligibility within a language

community and between speakers from the same region.

6.4 Conclusion

Throughout this thesis, the overarching issue in question was the acceptability of a

language for use in a disaster language policy. This approach highlighted the need

for language planning to ensure that accurate, intelligible, and trustable

communication be considered primary aspects to consider in the development of

policy. At the commencement of this thesis, the aim was to explore and identify how

the status quo from TWB (2017a) became the standard situation when processing

refugees en masse. The thesis reviewed and investigated the Ausbaucentric

ideologies enforced onto the field work contexts, whereby it is assumed that all

speakers of Arabic would be able to both understand and talk to each other in turn.

The reality is that inter-nationality communication between Arabic speakers is

hindered by linguistic variation, and an under-application of Abstand in the

demarcation process of what is a language, and whether a language variety is

acceptable for use in disaster management.

The consensus across disaster research (Jacobsen & Landau, 2003; Keselman et

al., 2005; Garnett & Moore, 2010) highlights that disaster policies should aim to be

objective and data-led, to reduce the potential impact of socio-political bias negatively impacting the effectiveness of alleviating the risk of harm to an at-risk population. Wright, (2016) shared a similar caveat, albeit related to general inter-cultural and inter-national language policies, stating that 'policy makers should be aware of the limitations of one's work in the wider context of a society' (pp. 12). This thesis supports this shared warning towards the development of disaster language policy, as presently there are established biases at work during the development and deployment of language support for refugees and aid workers. Whilst the socio-political aspects are important for a holistic assessment of a language, in the context of disasters, the focus needs to remain on providing information and communication which is, at the very least, intelligible for the listener to understand and to respond in kind.

In sum to respond to the titular question, the method to select a language for a disaster is to account for the veracity of a message and the speaker, to measure the Reach of using a selected language, but the first step that is needed going forward is the agreement as to what qualifies as a language variety, as a distinct language when Abstand, and only Abstand is used as the defining overarching criteria. Thus, in sociolinguistics, the first step advocated by this thesis is the further operationalisation of Abstand as an approach to language demarcation.

Appendix

Appendix A: Ethical approval for Intelligibility study (chapter 4)

PRIFYSGOL BANGOR UNIVERSITY

Ysgol Ieithoedd, Llenyddiaethau ac Ieithyddiaeth Prifysgol Bangor

School of Languages, Literatures and Linguistics Bangor University

Myfyriwr/*Student*: Conor Glackin (SLLL-041), 4 November 2019

Mae'r astudiaeth hon wedi cael ei chadarnhau o ran agweddau moesegol, yn dilyn ymgynghoriad gyda'r arolygwr (os perthnasol) a gyda'r swyddog Moeseg yr Ysgol. Mae rhyddid i'r fyfyriwr a enwir uchod barhau gyda chasglu'r data a gweithio ar yr astudiaeth.

This study has been approved with regards to ethical concerns, following consultation with the supervisor and the School Ethics officer. The student named above is now free to continue with collecting the data and working on the study.

Dr Thora Tenbrink
Swyddog Moeseg yr Ysgol / Darllenydd mewn Ieithyddiaeth Wybyddol
School Ethics officer / Reader in Cognitive Linguistics

Appendix B: Ethnical approval for trust study (chapter 5)

PRIFYSGOL
BANGOR
UNIVERSITY

Ysgol Ieithoedd, Llenyddiaethau, Ieithyddiaeth a'r Cyfryngau
Prifysgol Bangor

School of Languages, Literatures, Linguistics and Media
Bangor University

Ymchwilydd/*Researcher*: Conor Glackin
21/03/2021

Mae'r astudiaeth hon wedi cael ei chadarnhau o ran agweddau moesegol, yn dilyn ymgynghoriad gyda'r arolygwr (os perthnasol) a gyda'r swyddog Moeseg yr Ysgol. Mae rhyddid i'r fyfyriwr a enwir uchod barhau gyda chasglu'r data a gweithio ar yr astudiaeth.

This study has been approved with regards to ethical concerns, following consultation with the supervisor and the School Ethics officer. The student named above is now free to continue with collecting the data and working on the study.

Dr Dyfrig Jones
Swyddog Moeseg yr Ysgol
School Ethics officer

Appendix C: Sentence stimulus used in intelligibility study

English	Arabic	Variety
She made the bed with clean sheets.	غيرت شراشف السرير	Gulf
The boat sailed along the coast.	مشي القارب على الشاطئ	Gulf
The watchdog gave a warning growl.	نبح كلب الحراسة للتنبيه	Gulf
She wore a feather in her cap.	لبست قبعة بريش	Gulf
He caught the fish in his net.	علكت السمكة في الشبكة	Gulf
The airplane went into a dive.	نزلت الطائرة اضطراري	Gulf
Hold the baby on your lap.	امسك الطفل على حضنك	Gulf
On the beach we play in the sand.	لعبنا في الشاطئ بالرمل	Gulf
The lonely bird searched for its mate.	دور الطير على ونيس له	Gulf
They drank a whole bottle of gin.	شارب قارورة خمر كاملة	Gulf
After his bath he wore a robe.	لبس الروب بعد الحمام	Gulf
The soup was served in a bowl.	حط الشربة في الزبدية	Gulf
Lubricate the car with grease.	زيت العربية	Egyptian
The workers are digging a ditch.	العمال بيحفروا	Egyptian
They marched to the beat of the drum,	مشوا علي دق الطبل	Egyptian
Harry slept on the folding cot.	هاري نام علي سرير الاطفال اللي بيطبق	Egyptian
The cow was milked in the barn.	حلبوا البقرة في المزرعة	Egyptian
The cushion was filled with foam.	ملوا الخدادية بورق الفوم	Egyptian
The cookies were kept in a jar.	عانوا البسكوت في برطمان	Egyptian
The kitten climbed out on a limb.	القطة الصغيرة طلعت علي رجله	Egyptian
Get the bread and cut me a slice.	هات العيش و اقطعلي حتة	Egyptian
The burglar escaped with the loot.	الحرامي هرب باللي سرقه	Egyptian

This camera is out of film.	شريط الفيلم خلص في الكاميرا	Egyptian
Her hair was tied with a blue bow.	شعرها مربوط بتوكة فيونكة	Egyptian
The bird of peace is the dove.	الحمام هو طائر السلام	Egyptian
The car was parked at the curb.	كانت السيارة متوقفة على الرصيف.	MSA
We heard the ticking of the clock.	سمعنا دقات الساعة.	MSA
Instead of a fence, plant a hedge.	بدلاً من السياج ، قم بزرع سياج نباتي.	MSA
The firemen heard her frightened scream.	سمع رجال الإطفاء صراخها الخائف.	MSA
I ate a piece of chocolate fudge.	أكلت قطعة من حلوى الشوكولاتة.	MSA
To open the jar, twist the lid.	افتح الجرة ، لف الغطاء.	MSA
Paul took a bath in the tub.	استحم بول في حوض الاستحمام.	MSA
The candle burned with a bright flame.	تحترق الشمعة بلهب ساطع.	MSA
The chicken pecked corn with its beak.	الدجاج منقور الذرة مع منقاره.	MSA
The chicks followed the mother hen.	تتبع الكتاكيت أم الدجاجة.	MSA
We swam at the beach at high tide,	سبحنا على الشاطئ عند ارتفاع المد ،	MSA
My son has a dog for a pet,	ابني لديه كلب أليف ،	MSA

English	Arabic	Variety
go left at the fork past over the rail track turn east at the roundabout go north at the second junction then turn left	اتجه يسارًا عند مفترق الطرق مرورًا بمسار السكة الحديدية ثم انعطف شرقًا عند الدوار ثم اتجه شمالًا عند التقاطع الثاني ثم انعطف يسار	Gulf
go right at the fork straight over the junction turn north past the oval turn west at the junction then turn south at the roundabout	اتجه يمينًا عند مفترق الطرق وواصل السير الى بعد التقاطع ثم انعطف شمالًا بعد المنعطف ذات الشكل البيضاوي باتجاه الغرب عند التقاطع ثم انعطف جنوبًا عند الدوار	Gulf
go right at the fork turn left at the junction go straight over the junction continue around the corner past the junction	اتجه يمينًا عند مفترق الطرق ، ثم اتجه يسارًا عند التقاطع ، استمر في السير الى التقاطع واستمر حول الزاوية بعد التقاطع	Gulf
go left at the junction turn east at the roundabout turn right between the railways turn right past the oval turn south at the junction	اتجه نحو يسارًا عند التقاطع ثم انعطف شرقًا عند الدوار ثم انعطف يمينًا بين السكك الحديدية ثم انعطف يمينًا بعد الانعطاف البيضاوي ثم اتجه جنوبًا عند التقاطع	Gulf
go left at the junction turn right at the roundabout turn left past the semicircle continue to the roads end turn left	اتجه يسارًا عند التقاطع ثم انعطف يمينًا عند الدوار ثم انعطف يسارًا بعد نصف الدائرة وتابع السير إلى نهاية الطريق ثم انعطف يسارًا	Gulf
go right at the junction head north over the railway turn right past the railways head over the roundabout head right at the junction	اتجه يمينًا عند رأس التقاطع ثم اتجه شمالًا فوق السكة الحديدية ،ثم انعطف يمينًا بعد السكك الحديدية الرئيسية ثم اتجه يمينًا عند التقاطع الرئيسي	Gulf
go right at the fork turn north at the junction walk away from the rainbow go straight over the roundabout take the first left	خود يمينك عند مفرق الطرق لف شَمَال عند التفاطع إمشي بعيد عن مبنى القوس قزح امشي دغري و عدي الميدان بشكل مستقيم و خود اول شمال	Egyptian
go left at the fork straight over the railway continue north over the roundabout stay on the right at the junction turn left after the junction	خود شمالك عند مفرق الطرق و علي طول عدى المزلقان كمل شَمَال و عدي الميدان خليك يمين و عند التقاطع خود شمال بعد التقاطع	Egyptian
go right at the fork continue to the roads end then turn north to the junction turn left towards the rainbow then turn north at the junction	خود يمينك عند مفترق الطرق كمل لنهاية الطريق و بعدين لف شَمَال للتقاطع الأولاني و لف شمال ناحية مبنى القوس قزح و بعدين لف شَمَال عند التقاطع التاني	Egyptian
go left at the split walk right at the roundabout turn south at the roundabout head north after the railway turn left at the junction	خود شمال التفريعة إمشي يمين الميدان و لف ناحية الجنوب عند الميدان اللي بعديه و بعد كدا إتجه شَمَالَ بعد المزلقان و لف شمال عند التقاطع	Egyptian

go left at the split turn west past the diamond continue over the roundabout head north past the junction turn right	Egyptian	خود شمالك عند التفريعة و لف غرب مبنى الماسة كمل و عدى الميدان إتجه ناحية الشَّمَال و بعد التقاطع لف يمين
go right at the split walk north at the crossroads turn right at the junction continue east over the roundabout	Egyptian	خود يمين عند التفريعة إمشي ناحية الشَّمَال و عند تقاطع الشارعين خود يمينك و عند التقاطع اللي بعده كمل ناحية الشرق و إنت بتعدي الميدان
go left at the fork head over the railway go north at the roundabout turn right at the junction continue past the railway	Egyptian	خود شمالك عند مفترق الطرق روح ناحية المزلقان و عديه و إمشي ناحية الشَّمَال و عند الميدان لف يمين و عند التقاطع كمل بعد المزلقان.
go left at the fork turn west at the roundabout continue past the junction head west past the square then turn north	MSA	اتجه يسارًا عند مفترق الطرق ثم انعطف غربًا عند الدوار واستمر غرباً عند التقاطع ثم انعطف شمالًا..
go right at the fork turn north past the railway turn left at the junction turn right at the roundabout take the next left	MSA	اتجه يمينًا عند مفترق الطرق ثم انعطف شمالًا بعد خط السكة الحديد انعطف يسارًا عند التقاطع انعطف يمينًا عند الدوار وخذ اليسار التالي
go left at the fork turn east at the roundabout continue east at the crossroad turn north at the roundabout then turn south at the junction	MSA	اتجه يسارًا عند مفترق الطرق ثم انعطف باتجاه الشرق عند الدوار واستمر شرقاً عند مفترق الطرق ثم انعطف شمالاً عند الدوار ثم انعطف جنوباً عند التقاطع
go right at the junction walk south from the railway turn left until the roundabout turn west past both rail tracks then turn south	MSA	اتجه نحو اليمين عند التقاطع ، ثم اتجه جنوبًا عند السكة الحديدية ، ثم انعطف يسارًا عند الدوار ، انعطف غربًا مرورًا بخطي السكك الحديدية ، ثم انعطف جنوبًا
go left at the junction straight over the roundabout stay right at the junction continue past the roundabout turn north before the railway	MSA	اتجه نحو اليسار عند التقاطع واستمر السير فوق الدوار ابق يمينًا عند التقاطع واستمر بعد الدوران ، ثم انعطف شمالًا قبل السكة الحديد
go right at the junction walk left after the junction turn east over the roundabout turn left at the junction then east at the crossroad	MSA	اتجه نحو اليمين عند التقاطع ثم انعطف يسارًا بعد التقاطع ثم انعطف شرقًا فوق الدوار ثم انعطف يسارًا عند التقاطع ثم اتجه شرقًا عند مفترق الطرق

References

Abbas, F., Jalil, M. K., Zaki, H, N., Irfan, F. (2020), 'Implicit Measure of Language Attitude: Study of Punjabi Native Speakers by Using Matched Guise Technique'. *International Journal of Innovation, Creativity and Change*, 13(1), pp. 194-206.

Abdel-Rahman, M, S, A, F. (2016), *I Speak Better Arabic: Arabic Native Speakers' Perception of Arabic Dialects.*

Abdul-Mageed, M., Alhuzali, H., Elaraby, M. (2018), You tweet what you speak: A city-level dataset of arabic dialects. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2018).

Abelson, J, L., Erickson, T, M., Mayer, S, E., Crocker, J., Briggs, H., Lopez-Duran, N, L., Liberzon, I. (2014), Brief cognitive intervention can modulate neuroendocrine stress responses to the Trier Social Stress Test: Buffering effects of a compassionate goal orientation. *Psychoneuroendocrinology*, 44, pp. 60-70.

Aboagye Da-Costa, C., Adade-Yeboah, A. (2019), Language Practice and the Dilemma of a National Language Policy in Ghana: The Past, Present and Future. *International Journal of Humanities and Social Science* 9(3) pp. 142-149

Abrams, D., Hogg, M. A., Hinkle, S., & Often, S. (2005). The Social Identity Perspective on Small Groups. In M. S. Poole & A. B. Hollingshead (Eds.), *Theories of small groups: Interdisciplinary perspectives*, pp. 99–137 Sage Publications, Inc.

Abushariah, M.A., Ainon, R.N., Zainuddin, R., Elshafei, M., Khalifa, O.O. (2012), Phonetically rich and balanced text and speech corpora for Arabic language. Language resources and evaluation, 46(4), pp. 601-634.

African Union, (2022). Decision On The Use Of Kiswahili As An Au Working Language Of The African Union: ASSEMBLY OF THE UNION Thirty-Fifth Ordinary Session. Assembly/AU/Dec. 832(XXXV).

Ager, D. (2001), *Motivations in Language Planning and Language Policy.* London: Multilingual Matters.

Aghayev, E. (2013), Analysis and background of the "Arab Spring" in Libya. *European Researcher*, *39*(1-2), pp. 193-198.

Agustín, Ó, G., Jørgensen, M, B. (2018), *Solidarity and the 'refugee Crisis' in Europe*. Springer.

Akbacak, M., Vergyri, D., Stolcke, A., Scheffer, N., Mandal, A. (2011), Effective Arabic dialect classification using diverse phonotactic models. In *Twelfth Annual Conference of the International Speech Communication Association.* pp. 737-740.

Al Suwaiyan, L, A. (2018), Diglossia in the Arabic language. *International Journal of Language and Linguistics*, 5(3), pp. 228-238.

Alansary, S. and Nagi, M., (2014), The international corpus of Arabic: Compilation, analysis and evaluation. In *Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP)* pp. 8-17.

Al-Bazzaz, M, A., Ali, S, A,. (2020), A Sociolinguistic Study of Selected Iraqi-Arabic Dialects with Reference to TV Series. *Journal of Language Studies,* 3(3), pp. 68-85.

Albirini, A. (2011), The sociolinguistic functions of codeswitching between Standard Arabic and Dialectal Arabic. *Language in society*, 40(5), pp. 537-562.

Albirini, A. (2014), The role of the colloquial varieties in the acquisition of the standard variety: The case of Arabic heritage speakers. *Foreign Language Annals*, 47(3), pp. 447-463.

Albirini, A., Benmamoun, E. (2022). Arabic Diglossia and Heritage Arabic Speakers. In: Saiegh-Haddad, E., Laks, L., McBride, C. (eds) *Handbook of Literacy in Diglossia and in Dialectal Contexts*. Literacy Studies, vol 22. Springer

Albirini, A., Benmamoun, E., Saadah, E. (2011), Grammatical Features of Egyptian And Palestinian Arabic Heritage Speakers' oral Production. *Studies in Second Language Acquisition,* 33(2), pp. 273-303.

Albury, N., 2020. 18 Language attitudes and ideologies on linguistic diversity. In *Handbook of Home Language Maintenance and Development,* pp. 357-376 De Gruyter Mouton.

Al-Hajj S, Mokdad AH, Kazzi A, (2021), Beirut explosion aftermath: lessons and guidelines. Emergency *Medicine Journal* 2021; 38, pp. 938-939

Al-Jallad, A. (2019), The Classification of the Languages of North Arabia: Remarks on the Semitic Language Family Tree of the 2nd edition of Routledge's The Semitic Languages. *International Journal of Arabic Linguistics*, 5(2), pp. 86-99.

Al-Kahtany, A. H. (1997). The "Problem" of Diglossia in the Arab World: An Attitudinal Study of Modern Standard Arabic and the Arabic Dialects. *Al-'Arabiyya*, *30*, pp. 1–30.

Allen, A.P., Kennedy, P.J., Cryan, J.F., Dinan, T.G. Clarke, G. (2014) Biological and psychological markers of stress in humans: Focus on the Trier Social Stress Test. *Neuroscience & Biobehavioral Reviews*, 38, pp. 94-124.

Almujaiwel, S. (2020), An Overview of the Lexical Variation Between Arabic Lexicons and Natural Arabic Language. *International Journal of Arabic Linguistics*, 6(1-2), pp. 91-107.

Alresaini, S. (2016), Acquisition of Modern Standard Arabic by Speakers of Different Arabic Colloquial Varieties: Resumption in Object Relative Clauses. *Arab World English Journal (AWEJ)* 7(1) pp. 202-224.

Al-Rojaie, Y. (2013), Regional dialect leveling in Najdi Arabic: The case of the deaffrication of [k] in the Qaṣīmī dialect. *Language Variation and Change*, 25(1), pp. 43-63.

Alruily, M. (2020), Issues of dialectal Saudi twitter corpus. *Int. Arab J. Inf. Technol.*, 17(3), pp. 367-374.

Al-Sobh, M, A., Abu-Melhim, A, R, H., Bani-Hani, N, A. (2015), Diglossia as a result of language variation in Arabic: Possible solutions in light of language planning. *Journal of Language Teaching and Research*, 6(2), pp. 274.

Alves, M., Ode, C., Strömbergsson, S. (2020), Dealing with the unknown–addressing challenges in evaluating unintelligible speech. *Clinical Linguistics & Phonetics*, 34(1-2), pp. 169-184.

Al-Wash, L. (2016), Language Varieties & the Effect of Geography "Standard Arabic & Arabic Dialectal Languages as an example". *Scientific Journal of Faculty of Education* 2(6) pp. 25-29

Amara, M., 2018. Challenges of Arabic language education policies in the Arab World. In *The Routledge Handbook of Arabic Linguistics*, pp. 546-559.

Amayreh, M. (2003), Completion of the Consonant Inventory of Arabic. *Journal of Speech, Language, and Hearing Research*, 46, pp. 517–529.

Amayreh, M., Dyson, A. (1998), The acquisition of Arabic Consonants. *Journal of Speech, Language, and Hearing Research*, 41, pp. 642–653

Ammon, U. (2006), Language conflicts in the European Union; On finding a politically acceptable and practicable solution for EU institutions that satisfies diverging interests. *International Journal of Applied Linguistics*. 16(3) pp. 319-338

Anand, S., Stepp, C, E. (2015), Listener perception of monopitch, naturalness, and intelligibility for speakers with Parkinson's disease. *Journal of Speech, Language, and Hearing Research,* 58(4), pp. 1134-1144.

Androutsopoulos, J. (2014), *Mediatization and Sociolinguistic Change*. Walter de Gruyter, pp. 5-7

Anwyl-Irvine, A.L., Massonnié, J., Flitton, A., Kirkham, N., Evershed, J, K. (2020), 'Gorilla in our midst: an online behavioural experiment builder.' *Behavior Research Methods*. 52, pp. 388–407

Arai, K. (2013a), The Communication of Disaster Prevention and Relevance - How to Transmit Tsunami Warnings Effectively. *Keieironshu* No. 81: 91–105.

Arai, K. (2013b), How to transmit disaster information effectively: A linguistic perspective on Japan's Tsunami Warnings and Evacuation Instructions. *International journal of disaster risk science*, 4(3), pp. 150-158.

Arora, S., Sevdalis, N., Nestel, D., Woloshynowych, M., Darzi, A., Kneebone, RL. (2010), The impact of stress on surgical performance: a systematic review of the literature. *Surgery*. 147, pp. 318–30.

Asadi, I, A., Abu-Rabia, S. (2021), The Impact of Diglossia on Phonological Processing. *Reading Psychology,* 42(7), pp. 685-699.

Ashcraft, M, H. (2019), 'Models of math anxiety', In Irene C. M, Sara C., Ann D.(eds.) *Mathematics Anxiety.* Oxon, UK, pp. 1-19.

Atkinson, R, C., and Shiffrin, R, M. (1968), Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* 2(1) pp. 89-195

Audacity Team. (2008), Available at http://audacityteam.org/ (Accessed 24/07/22)

Austermuhl, F. (2014), *Electronic tools for translators*. Routledge.

Avancini, C., Szűcs D. (2019), 'Psychophysiological correlates of mathematics anxiety', In Mammerella, I., Caviola, S., Dowker, A., (eds.) *Mathematics Anxiety.* Oxon, UK, pp. 42-61.

Bager, L., Hansen, K.S., Andersen, C.J. Wang, S.J., 2018. Does multidisciplinary rehabilitation of tortured refugees represent 'value-for-money'? A follow-up of a Danish case-study. *BMC health services research*, *18*(1), pp. 1-14.

Bagilishya, D. (2000), Mourning and Recovery from Trauma: In Rwanda, Tears Flow Within. *Transcultural Psychiatry*, 37(3), pp. 337-353.

Bakker, E., Singleton, M. (2016), Foreign Fighters in the Syria and Iraq Conflict: Statistics and Characteristics of a Rapidly Growing Phenomenon. In: A. de Guttry, F. Capone & P. Christophe., eds. *Foreign Fighters under International Law and Beyond*. The Hague: T.M.C. Asser Press, pp. 9-25.

Baldauf Jr, R, B., Kaplan, R, B. (2005), Language-in-education policy and planning. *Handbook of research in second language teaching and learning*, pp. 1037-1058.

Balkan, B., Tumen, S. (2016), Immigration and prices: quasi-experimental evidence from Syrian refugees in Turkey. *Journal of Population Economics*, 29(3), pp. 657-686.

Bardel, C., Falk, Y. and Lindqvist, C., 2013. 10 Multilingualism in Sweden. *Current Multilingualism*, pp. 247-270..

Barnett, M. (2003), *Eyewitness to a Genocide: The United Nations and Rwanda.* New York: Cornell University Press.

Bassiouney, R. (2009), *Arabic sociolinguistics*. Edinburgh: Edinburgh University Press.

Bassiouney, R. (2014), *Language and identity in modern Egypt*. Edinburgh University Press.

Bassnett, S. (2013), *Translation studies*. Routledge, UK.

Bean, H., Sutton, J., Liu, B, F., Madden, S., Wood, M, M., Mileti, D, S. (2015), The study of mobile public warning messages: A research review and agenda. *Review of Communication*, 15(1), pp. 60-80.

Benmamoun, E., Bassouiney, R. (eds) *The Routledge Handbook of Arabic Linguistics*/. Routledge, UK, pp. 48-61

Bentley, R, S. (2015), "Conditional Sentences in Egyptian Colloquial and Modern Standard Arabic: A Corpus Study. All Theses and Dissertations. 4440. Available at: https://scholarsarchive.byu.edu/etd/4440 (Accessed 24/07/2022)

Berry, M., Garcia-Blanco, I., Moore, K. (2016), *Press coverage of the refugee and migrant crisis in the EU: A content analysis of five European countries*.

Besançon, M, L. (2005), Relative Resources: Inequality in Ethnic Wars, Revolutions, and Genocides. *Journal of Peace Research*, 42(4), pp. 393-415.

Bhatia, T, K. (2018), Accent, intelligibility, mental health, and trauma. *World Englishes*, 37(3), pp. 421-431.

Bidaoui, A. (2016), Discourse markers of elaboration in Maghrebi and Egyptian dialects: A socio-pragmatic perspective. *International Journal of Arabic Linguistics*, 2(1), pp. 19-45.

Bidaoui, A. (2020), Attitudes toward Arabic and Foreign Elaboration DMs in Three Dialects of Arabic. Al-ʿArabiyya: *Journal of the American Association of Teachers of Arabic*, 53, pp. 57-83.

Birkett, M, A. (2011), The Trier Social Stress Test protocol for inducing psychological stress. *JoVE (Journal of Visualized Experiments),* 56 (e3238), pp. 1-6.

Bishop, V, M, D. (2003), *Test for Reception of Grammar: Version 2 ; TROG-2.* Harcourt Assessment, New York.

Bolin, B. (2007), Race, Class, Ethnicity, and Disaster Vulnerability. In Anserson, W, A. Kennedy, P,J,. Kessler, E (eds.) *Handbook of Disaster Research. Handbooks of Sociology and Social Research*. Springer, New York, pp. 113-129

Bong, C.L., Fraser, K., Oriot, D. (2016), Cognitive load and stress in simulation. In Comprehensive healthcare simulation: *Pediatrics*. pp. 3-17

Bouamor, H., Habash, N., and Oflazer, K. (2014), A multidialectal parallel corpus of Arabic. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* pp. 1240-1245.

Boucsein, W. (1992), *Electrodermal activity*. New York: Plenum Press

Bourantonis, D. (2004), *The History and Politics of UN Security Council Reform*. London: Routledge.

Bouwer, L. (2007), Intercomprehension and mutual intelligibility among southern Malagasy languages. *Language Matters*, 38(2), pp. 253-274.

Bowden, J. (2019). Defending the Heartland: Saudi Arabia's Response to the Arab Spring. In: Çakmak, C., Özçelik, A. (eds) *The World Community and the Arab Spring*. Palgrave Macmillan, Cham.

Boyarsky, I., Shneiderman, A., 2002. Natural and hybrid disasters—Causes, effects, and management. *Advanced Emergency Nursing Journal*, 24(3), pp. 1-25.

British Council. (2022), *What is the IELTS?* Available at: https://takeielts.britishcouncil.org/take-ielts/what-ielts. (Accessed 24/07/22)

Britton, N, R. (2007), National Planning and Response: National Systems. In Anserson, W, A. Kennedy, P, J., and Kessler, E (eds.) *Handbook of Disaster Research. Handbooks of Sociology and Social Research*. Springer, New York, pp. 347-367

Broadbent, D, E. (1971), *Decision and stress*. Academic Press.

Broussard, G., Rubenstein, L.S., Robinson, C., Maziak, W., Gilbert, S, Z., DeCamp, M. (2019), Challenges to ethical obligations and humanitarian principles in conflict settings: a systematic review. *Journal of International Humanitarian Action*, 4(1), pp. 1-13.

Brustad, K., 2017. Diglossia as ideology. In Hoigilt, J., Mejdell, G.(eds.) *The politics of written language in the Arab world* pp. 41-67.

Butler, N., Stoyanova Russell, D. (2018), No funny business: Precarious work and emotional labour in stand-up comedy. *Human Relations*, *71*(12), pp. 1666-1686.

Cabrera, M, P, M. (2014), Textbook use training in EFL teacher education. *Utrecht studies in language communication*. 27 pp. 267-286

Cadora, F, J. (1989), Linguistic Change and the Bedouin-Sedentary Dichotomy in Old Arabic Dialects. Anthropological Linguistics, pp. 264-284.

Cadwell, P., 2021. Translation and interpreting in disaster situations. *The Routledge Handbook of Translation and Health*, pp. 253-268.

Calhoun, C. (2008), A World of Emergencies: Fear, Intervention, and the Limits of Cosmopolitan Order. *Canadian Review of Sociology/Revue canadienne de sociologie*, 41(4), pp. 373-395.

Carter, W, N. (2008), *Disaster management: A disaster manager's handbook*.

Cenoz, J., Gorter, D. (2017), Minority languages and sustainable translanguaging: threat or opportunity? *Journal of Multilingual and Multicultural Development*, 38(10), pp. 901–912.

Centre for Disease Control and Prevention. (2002), *Crisis Emergency Risk Communication*. 2002 Edition.

Centre for Disease Control and Prevention. (2014), *Crisis Emergency Risk Communication*. 2014 Edition.

Chakrani, B. (2015), Arabic interdialectal encounters: Investigating the influence of attitudes on language accommodation. *Language & Communication*, 41, pp. 17-27.

Chakrani, B., (2020), Languages in Conflict: Examining the Status of Standard Arabic and French in Morocco. *Al-ʿArabiyya: Journal of the American Association of Teachers of Arabic 53*, pp. 37-56.

Chejne, A.G. (1968), *The Arabic language: Its role in history*. U of Minnesota Press.

Cheng, C, C. (1996), Quantifying dialect mutual intelligibility. *In New horizons in Chinese linguistics*, pp. 269-292

Clarke, A, J., Isphording, I, E. (2015), Language barriers and immigrant health production. IZA Discussion No. 8846.

Clarke, K., (2018), When do the dispossessed protest? Informal leadership and mobilization in Syrian refugee camps. *Perspectives on Politics*, *16*(3), pp. 617-633.

Clopper, C, G., Bradlow, A, R. (2008), Perception of dialect variation in noise: Intelligibility and classification. Language and speech, 51(3), pp. 175-198.

Clover, K., Lambert, S.D., Oldmeadow, C., Britton, B., King, M.T., Mitchell, A.J. and Carter, G.L., (2020). Apples to apples? Comparison of the measurement properties of hospital anxiety and depression-anxiety subscale (HADS-A), depression, anxiety and stress scale-anxiety subscale (DASS-A), and generalised anxiety disorder (GAD-7) scale in an oncology setting using Rasch analysis and diagnostic accuracy statistics. *Current Psychology*, pp. 1-10.

Commission of the European Communities. (2008), Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Multilingualism: an asset for Europe and a shared commitment. SEC(2008) 2443.

Conway, A., Kane, M., Bunting, M., Hambrick, D., Wilhelm, O., and Engle, R. (2005), Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), pp. 769-786.

Coombs, W, T. (2007), Protecting organization reputations during a crisis: The development and application of situational crisis communication theory. *Corporate reputation review*, 10(3), pp. 163-176

Cooper, T. (2013), *Great Lakes Conflagration: Second Congo War, 1998 2003*. Helion and Company.

Cortina, J, M. (1993), What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, *78*(1), pp. 98.

Cotter, W, M. (2017), Not-so-strange bedfellows: Documentation, description, and sociolinguistics in Gaza. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 62(4), pp. 596-613.

Council of Europe (1995) FRAMEWORK CONVENTION FOR THE PROTECTION OF NATIONAL MINORITIES AND EXPLANATORY REPORT, H (95)10. Available at: https://rm.coe.int/16800c10cf (Accessed 24/07/2020)

Council of Europe. (1992), *European Charter for Regional or Minority Languages, ETS 148.*

Council of Europe. (1995), Framework Convention for the Protection of National Minorities

Council of Europe. (2010), The European Charter for Regional Or Minority Languages: Collected Texts. Council of Europe Publishing, Strasbourg. pp. 9-28

Cowan, N. (2008), What are the differences between long-term, short-term, and working memory?. *Progress in brain research*, 169, pp. 323-338.

Cronbach, L, J. (1951), Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3)

Crystal, D. (2003), *English as a global language.* Cambridge university press.

d'Almeida, A, C., Otcu-Grillman, B. (2013), The Portuguese language in the United Nations–framing policy design. *International journal of the sociology of language*, 2013(224), pp. 1-23.

Dankert, A, C. (2017), Europe under pressure: the development of the European Union under the influence of the Arab Spring, the refugee crisis and the global threat of terrorism. Tectum Wissenschaftsverlag.

de Azevedo, C, V. (2020), ISIS Resurgence in Al Hawl Camp and Human Smuggling Enterprises in Syria. *Perspectives on terrorism*, *14*(4), pp. 43-63.

De Jong, K., and Zawaydeh, B. (2002), Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, *30*(1), pp. 53-75.

De Jong, R. (2004), Variation and change in Arabic urban vernaculars. In Haak, M,. De Jong, R., Versteegh, K., *Approaches to Arabic Dialects*, Brill, Leiden. pp. 177-206

De Varennes, F. (2021), *Language, minorities and human rights.*

Demuth, J, L. (2002), *Countering terrorism: Lessons learned from natural and technological disasters*. Washington, DC: National Academy of Sciences.

Derwing, T, M., Munro, M, J. (1997), Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in second language acquisition*, *19*(1), pp. 1-16.

Dimmendaal, G, J. (2008), Language ecology and linguistic diversity on the African continent. *Language and linguistics compass*, 2(5), pp. 840-858.

Dixon, T. L., Linz, D. (2002). Television news, prejudicial pretrial publicity, and the depiction of race. *Journal of Broadcasting & Electronic Media*, 46, pp. 112-136

Djennane, T. (2014), Diglossia's Stability in the Arab World: Algeria as an Instance. *IOSR Journal of Humanities and Social Science*, 19(11), pp. 52-56.

Doğan, A. (2016), Anybody Down There? Emergency and Disaster Interpreting in Turkey. In *Mediating Emergencies and Conflicts*, pp. 59-83 Palgrave Macmillan.

Dong, Y., Gui, T., Zhang, M., and Zhong, Y. (2022), Language Confidence and Acquisition: Perception of Accent Affects Oral Speaking Skill in Second Language Acquisition. In *2021 International Conference on Public Art and Human Development (ICPAHD 2021)*, pp. 1123-1126

Donner, W, R., Rodriguez, H., Diaz, W. (2007), Public warning response following tornadoes in New Orleans, LA, and Springfield, MO: A sociological analysis. In *Second Symposium on Policy and Socio-economic Research, 87th Annual Meeting of the American Meteorological Society,* San Antonio, Texas.

Dowker, A. (2019), 'Mathematics anxiety and performance', In Irene C. M, Sara C. and Ann D.(ed.) *Mathematics Anxiety*, pp. 62-76.

Dryer, M, S., Haspelmath, M. (2013), *The world atlas of language structures online.*

Dückers, M, L. (2017), Anticipating the Psychosocial Impact of Disasters and Crises: The Need for an Interdisciplinary Social Science Framework. *Prehospital and Disaster Medicine*, *32*(S1), pp. S181-S182.

Eberhard, D, M., Simons, G, F., Fennig, C, D. (2022), Ethnologue: Languages of the World. Twenty-fifth edition. Dallas, Texas: SIL International.

Edgington, D.W. (2021), Planning for Earthquakes and Tsunamis: Lessons from Japan for British Columbia, Canada. *Progress in Planning*, pp. 100-626.

Edwards, R., 1998. A critical examination of the use of interpreters in the qualitative research process. *Journal of ethnic and migration studies*, *24*(1), pp. 197-208.

El Zarka, D., Hellmuth, S. (2009), Variation in the intonation of Egyptian Formal and Colloquial Arabic. *Langues et Linguistique*, 22, pp. 73-92.

Elmahdy, M., Gruhn, R., Minker, W., Abdennadher, S. (2009), Survey on common Arabic language forms from a speech recognition point of view. In proceeding of International conference on Acoustics (NAG-DAGA), pp. 63-66

Embarki, M., Yeou, M., Guilleminot, C., Al Maqtari, S. (2007), An acoustic study of coarticulation in modern standard Arabic and dialectal Arabic: pharyngealized vs. nonpharyngealized articulation. ICPhS XVI, pp. 141-146.

Engle, R., Tuholski, S., Laughlin, J., Conway, A. (1999), Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology* 128(3), pp. 309-331.

Etkin, D. (2014), *Disaster theory: an interdisciplinary approach to concepts and causes*. Butterworth-Heinemann.

European Union. (2007), *Consolidated version of the Treaty on the Functioning of the European Union*, 2008/C 115/01

European Union. (2008), Council Resolution of 21 November 2008 on a European strategy for multilingualism. *Official Journal of the European Union*. 320, pp. 1-3.

European Union. (2010), Charter of Fundamental Rights of the European Union. Official Journal of the European Union C83, 53, European Union, pp. 380

European Union. (2014), Council conclusions of 20 May 2014 on multilingualism and the development of language competences. 2014/C183/06, pp. 26-29

Fanning, R, M., Gaba, D, M. (2007), The role of debriefing in simulation-based learning. *Simulation in healthcare*, 2(2), pp. 115-125.

Farazmand, A. (2001), *Handbook of crisis and emergency management.* CRC Press.

Farghaly, A. (2010), The Arabic language, Arabic linguistics and Arabic computational linguistics. *Arabic computational linguistics*, pp. 43-81.

Farghaly, A., Shaalan, K. (2009), Arabic natural language processing: Challenges and solutions. *ACM Trans. Asian Lang. Inform. Process*. 8,4(14) pp. 1-22

Federici, F.M., (2016). *Mediating emergencies and conflicts*. Houndmills: Palgrave Macmillan.

Ferguson, C, A. (1959), Diglossia. *WORD*, 15(2), pp. 325-340.

Ferguson, C, A. (1991), Diglossia revisited. *Southwest Journal of Linguistics*, 10(1), pp. 214-234.

Ferguson, G. (2006), *Language planning and education*, Edinburgh: Edinburgh University Press.

Fiedler, S., Keller, C., Hanulíková, A. (2019), Social expectations and intelligibility of Arabic-accented speech in noise. In *Proceedings of the 19th International Congress of Phonetic Sciences*, Australia pp. 3085-3089.

Fierros, C., Jaloway, K., Jones, A., Mason, M, M., Shtepani, C., Shuler, M., Spiva, R., Ward, J., Weir, C. (2017), *Do NGOs Deliver?* The Role of NGOs in Responding to the Syrian Refugee Crisis in Jordan.

Fisanick, C. (2004), *The Rwanda genocide*. London: Greenhaven Press.

Fishman, J, A. (1967), Bilingualism with and Without Diglossia; Diglossia With and Without Bilingualism. *Journal of Social Issues*, 23, pp. 29-38

Fishman, J, A. (1991), *Reversing language shift*, Clevedon, UK, Multilingual Matters Ltd

Fitzpatrick, M, P. (2018), Colonialism, Postcolonialism, and Decolonization. *Central European History,* 51(1), pp. 83-89.

Fleisch, A. (2009), Language History in SE Angola*. The Ngangela-Nyemba dialect cluster. Language contact, language change, and history based on language sources in Africa*, pp. 97-111.

Flipsen Jr, P. (2006), Measuring the intelligibility of conversational speech in children. *Clinical linguistics & phonetics,* 20(4), pp. 303-312.

Fokaefs, A., Sapountzaki, K. (2021), Crisis Communication after Earthquakes in Greece and Japan: Effects on Seismic Disaster Management. *Sustainability*, 13(16), pp. 9257.

Frisch, J, U., Häusser, J, A., Mojzisch, A. (2015), The Trier Social Stress Test as a paradigm to study how people respond to threat in social interactions. *Frontiers in psychology*, 6(14), pp. 1-15

Frishkopf, L. (2010), Music and Media in the Arab World and Music and Media in the Arab World as Music and Media in the Arab World: A metadiscourse. In Firchkopt (eds.) *Music and Media in the Arab World*. The American University in Cairo Press pp. 1-67

Fukami, M., Hisamoto, N. (2010), A general view of Japanese disaster prevention system. *京都大学大学院経済学研究科 Working Paper*, (115).

Fuse, A., Navichkova, Y., Alloggio, K. (2018), Perception of intelligibility and qualities of non-native accented speakers. *Journal of Communication Disorders,* 71, pp. 37-51.

Gadalla, H, A. (2000b), *Comparative morphology of standard and Egyptian Arabic* (Vol. 5).

Garavan, H. (1998), Serial attention within working memory. *Memory & cognition*, 26(2), pp. 263-276.

Gardner-Chloros, P., Gogonas, N., McEntee-Atalianis, L. (2016), How can linguists contribute to the refugee crisis? Issues and responses. Public Event and Conference (Migrants and Language (s), *Éducation et sociétés plurilingues,* (40), pp. 95-101.

Garnett, J., Moore, M. (2010), Enhancing Disaster Recovery. *Journal of Homeland Security and Emergency Management*, 7(1),

Gass, S., Varonis, E, M. (1984), The effect of familiarity on the comprehensibility of nonnative speech. *Language learning*, 34(1), pp. 65-87.

Geiger, M., Pécoud, A. (2010), The politics of international migration management. In *The politics of international migration management*, pp. 1-20 Palgrave Macmillan, London.

Gerver, M. (2021a), Moving in an Unjust World, *Analysis*, 81(3), pp. 577–586.

Gerver, M. (2021b), Helping Refugees Where They Are. *Ethics & International Affairs*, 35(4), pp. 563-580.

Ghazali, S., Hamdi, R., Barkat, M. (2002), Speech rhythm variation in Arabic dialects. *In Speech Prosody 2002, International Conference.*

Ginsburgh, V., & Moreno-Ternero, J. D. (2022), Brexit and multilingualism in the European Union. *Metroeconomica*, 73, pp. 708– 731

Gomes, J, F. (2020), The health costs of ethnic distance: evidence from Sub-Saharan Africa. *Journal of Economic Growth*, *25*(2), pp. 195-226.

Gooskens, C. (2007), The contribution of linguistic factors to the intelligibility of closely related languages. *Journal of Multilingual and multicultural development*, 28(6), pp. 445-467.

Gooskens, C. (2017), *Dialect intelligibility. The handbook of dialectology*, pp. 204-218.

Gooskens, C. Van Heuven, V, J. (2021), *Mutual intelligibility. Similar languages, varieties, and dialects: A computational perspective*, pp. 51-95.

Gooskens, C., Swarte, F. (2017), Linguistic and extra-linguistic predictors of mutual intelligibility between Germanic languages. *Nordic Journal of Linguistics*, 40(2), pp. 123-147.

Gooskens, C., Van Bezooijen, R., and Van Heuven, V, J. (2015), *Mutual intelligibility of Dutch-German cognates by children: The devil is in the detail.* Linguistics, 53(2), pp. 255-283.

Gooskens, C., Van Heuven, V, J. (2020), How well can intelligibility of closely related languages in Europe be predicted by linguistic and non-linguistic variables?. *Linguistic approaches to bilingualism*, 10(3), pp. 351-379.

Gooskens, C., Van Heuven, V, J., Van Bezooijen, R., and Pacilly, J, J. (2010), Is spoken Danish less intelligible than Swedish?. *Speech Communication*, 52(11-12), pp. 1022-1037.

Gorilla Experimental Builder. (2020), Available at: www.gorilla.sc (Accessed 24/07/2022)

Gorter, D., Cenoz, J. (2017). Language education policy and multilingual assessment. *Language and Education*, 31(3), pp. 231–248.

Guerini, F. (2008), Multilingualism and language attitudes in Ghana: A preliminary survey. *Ethnorema*, 4(3).

Hachimi, A. (2013), The Maghreb-Mashriq language ideology and the politics of identity in a globalized Arab world. *Journal of Sociolinguistics*, 17(3), pp. 269-296.

Halliday, M. (2001), New ways of meaning: the challenge to applied linguistics. in: A. Fill and P. Mühlhäusler (eds), *The Ecolinguistics Reader: Language, Ecology and Environment*. London: Continuum.

Hammarström, H. (2019), *An inventory of Bantu languages*. In *The Bantu languages,* pp. 17-78 Routledge.

Hammond, A., (2007), *Popular culture in the Arab world: Arts, politics, and the media*. American Univ in Cairo Press.

Hart, R. (2009), Child refugees, trauma and education: interactionist considerations on social and emotional needs and development. *Educational Psychology in Practice*, 25(4), pp. 351-368.

Hartwright, C, E., Looi, C, Y., Sella, F., Inuggi, A., Santos, F, H., González-Salinas, C., García Santos, J, M., Kadosh, R, C., Fuentes, L, J. (2018), The neurocognitive architecture of individual differences in math anxiety in typical children. *Scientific reports*, 8(1), pp. 1-10.

Hassan, M.E., (2021), Reshaping Social Practice in Post–Arab Spring Egypt: Expression of Identity and Affiliation in New Media. *Cultural Production and Social Movements After the Arab Spring: Nationalism, Politics, and Transnational Identity*, pp.151-162.

Heaton, H. (2018), Media influence on implicit and explicit language attitudes (Doctoral dissertation).

Hellier, E., Edworthy, J., Weedon, B., Walters, K., Adams, A. (2002), The perceived urgency of speech warnings: Semantics versus acoustics. *Human Factors*, *44*(1), pp. 1-17.

Hewstone, M., Giles, H. (1997), Social groups and social stereotypes. *Sociolinguistics*, pp. 270-283.

Hintjens, H, M. (1999), Explaining the 1994 genocide in Rwanda. *The Journal of Modern African Studies*, 37(2), pp. 241-286.

Hodoshima, N. (2019), Effects of Urgent Speech and Congruent/Incongruent Text on Speech Intelligibility in Noise and Reverberation. *INTERSPEECH* pp. 3113-3117.

Hogg, M, A. (2016), Social identity theory. In *Understanding peace and conflict through social identity theory*, pp. 3-17 Springer, Cham.

Høigilt, J., Mejdell, G. (2017), The politics of written language in the Arab world: *Writing change* pp. 330

Holes, C. (1990), *Gulf Arabic*. Georgetown University Press, USA.

Holes, C. (2003), Quadriliteral verbs in the Arabic dialects of Eastern Arabia. Approaches to Arabic dialects: a collection of articles presented to Manfred Woidich on the occasion of his 60th birthday, pp. 97-116.

Holes, C. (2005), Form X of the verb in the Arabic dialects of Eastern Arabia. In Semitic Studies in Honour of Edward Ullendorff , pp. 115-125

Holes, C. (2006), The Arabic dialects of Arabia. In Proceedings of the seminar for Arabian studies. *Archaeopress*. pp. 25-34.

Holes, C. (2011), Language and Identity in the Arabian Gulf, *Journal of Arabian Studies*, 1(2), pp. 129-145

Holmes, J, S., De Piñeres, S, A, G. (2011), Conflict-induced displacement and violence in Colombia. *Studies in Conflict & Terrorism*, *34*(7), pp. 572-586.

Horesh, U., Cotter, W, M. (2016), Current research on linguistic variation in the Arabic-speaking world. Language and Linguistics Compass, 10(8), pp. 370-381.

Hornsby, B, W. (2004), The Speech Intelligibility Index: What is it and what's it good for?. *The Hearing Journal*, 57(10), pp. 10-17.

House, J. (2014), Translation quality assessment: Past and present. In House, J. (eds) *Translation: A multidisciplinary approach.* Palgrave Macmillan, London. pp. 241-264.

Hult, F, M. (2012), English as a Transcultural Language in Swedish Policy and Practice. *TESOL Quarterly: A Journal for Teachers of English to Speakers of Other Languages and of Standard English as a Second Dialect*, 46(2), pp. 230–257.

Hult, F, M., Hornberger, N, H. (2016), Revisiting orientations in language planning: Problem, right, and resource as an analytical heuristic. *The Bilingual Review/La Revista Bilingüe*, 33(3), pp. 30-49.

Huttunen, K., Keränen, H., Väyrynen, E., Pääkkönen, R., Leino, T. (2011), Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. *Applied ergonomics*, 42(2), pp. 348-357.

Hyatt, D. (2013), Stakeholders' perceptions of IELTS as an entry requirement for higher education in the UK. *Journal of Further and higher Education*, 37(6), pp. 844-863

Ibrahim, Z. (2021), A reflection on the relationship between Diglossia and its Effects on Modern Standard Arabic Lexical Variation and the Mental Lexicon. *Academia Letters*, pp. 2-8.

IELTS. (2022), *Home of the IELTS English Language Test* Available at: http://www.ielts.org (accessed: 24/07/22)

Imbo, I., Vandierendonck, A. (2008), Effects of problem size, operation, and working-memory span on simple-arithmetic strategies: differences between children and adults?. *Psychological research*, 72(3), pp. 331-346.

Imbo, I., Vandierendonck, A., Vergauwe, E. (2007), The role of working memory in carrying and borrowing. *Psychological research*, 71(4), pp. 467-483.

Inoue, H. (2012), Can We Use "Imperatives" in the Evacuation Instructions? - From the Viewpoint of the Effective Expressions for the Listeners. Broadcast Study and Survey, *NHK Culture and Broadcast Institute*, pp. 2–15

Isaksson, B. (1996), Arabic Dialectology: The State of the Art: Review of Dialectologia Arabica: A Collection of Articles in Honour of the Sixtieth Birthday of Professor Heikki Palva. *Orientalia Suecana*, *43*, pp. 115-132.

Issaev, L., Khokhlova, A., Korotayev, A. (2022). The Arab Spring in Yemen. In: Goldstone, J.A., Grinin, L., Korotayev, A. (eds) *Handbook of Revolutions in the 21st Century*. Societies and Political Orders in Transition. Springer, Cham.

Ivana, K., (2022), *Multilingualism: The language of the European Union*, EPRS: European Parliamentary Research Service

Ivanova, O. (2013), Bilingualism in Ukraine: defining attitudes to Ukrainian and Russian through geographical and generational variations in language practices. *Sociolinguistic studies*, 7(3), pp. 249-272.

Jacobsen, K., Landau, L, B. (2003), The Dual Imperative in Refugee Research: Some Methodological and Ethical Considerations in Social Science Research on Forced Migration. *Disasters*, 27(3), pp. 185-206.

Jastram, K., Newland, K. (2003), Family unity and refugee protection, in Feller, V. Türk, and F. Nicholson (eds) . Refugee Protection In International Law: UNHCR'S Global Consultations On International Protection. Cambridge University Press, UK,, pp. 555-603.

Johnsen, G, E., Kanagaratnam, P., Nordling, T., Wilberg, C., Asbjørnsen, A, E. (2017), Psychophysiological responses to an executive task in refugees with chronic PTSD and comorbid depression. Nordic psychology, 69(3), pp. 177-194.

Johnson, D, C. (2013), What is language policy?. In Johnson, D.C. (eds) *Language policy : Research and Practice in Applied Linguistics*. pp. 3-25. Palgrave Macmillan, London.

Jones, I., Askew, L. (2016) *Meeting the language challenges of NATO operations: Policy, practice and professionalization.* Springer.

Judge, A. (2007), *Linguistic policies and the survival of regional languages in France and Britain.* Palgrave Macmillan.

Juffermans, K. (2015), *Local languaging, literacy and multilingualism in a West African society.* Multilingual Matters.

Kabranian-Melkonian, S. (2015), Ethical concerns with refugee research. *Journal of Human Behavior in the Social Environment*, 25(7), pp. 714-722.

Kalikow, D, N., Stevens, K, N., Elliott, L, L. (1977), Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the acoustical society of America*, 61(5), pp. 1337-1351.

Kaplan, R, B. (2018), Is English Global? In Lian, P.C.G, Chua, C, Taylor-Leech, K., Williams, C. (eds) *Un(intended) language planning in a globalising world: Mutliple levels of players at work*. De Gruyter, Warsaw. pp. 15-38

Kaplan, R, B., Baldauf, R, B. (1997), *Language planning from practice to theory* (Vol. 108). Multilingual Matters.

Kapović, M. (2011), Language, ideology and politics in Croatia. *Slavia centralis*, 4(2), pp. 45-56.

Karnieli, A., Shtein, A., Panov, N., Weisbrod, N., Tal, A. (2019), Was drought really the trigger behind the Syrian civil war in 2011. *Water*, 11(8), pp. 15-64.

Kates, J, M., Arehart, K, H. (2005), Coherence and the speech intelligibility index. *The journal of the acoustical society of America,* 117(4), pp. 2224-2237.

Kaye, A, S. (2001), Diglossia: the state of the art. *International Journal of the Socioloigy of Language*, 2001(152), pp. 117-129

Kazeem, A, S., Suleiman, M, I. (2020), Prospects and Problems of Language Policy. *Britain International of Linguistics Arts and Education (BIoLAE) Journal*, 2(1), pp. 330-337

Kemper, N, S., Campbell, D, S., Earleywine, M., Newheiser, A. (2020), Likert, slider, or text? reassurances about response format effects, *Addiction Research & Theory*, 28:5, pp. 406-414

Keselman, A., Slaughter, L., Patel, V, L. (2005), Toward a framework for understanding lay public's comprehension of disaster and bioterrorism information. *Journal of Biomedical Informatics,* 38(4), pp. 331-344.

Keysar, B., Barr, D, J., Balin, J, A., Brauner, J, S. (2000), Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(1), pp. 32-38.

Khalifa, S., Habash, N., Abdulrahim, D., Hassan, S. (2016), *A Large Scale Corpus of Gulf Arabic*, Cornell University Library, arXiv.org, Ithaca.

Khalil, S. (2012), The evolution of the Arabic language through online writing: The explosion of 2011. In BRISMES 2012 Graduate Conference Papers: *Change and Continuity in the Arab World*.

Khepera, K., 2020. Language policy and decolonizing African education. *Journal of Management Policy and Practice*, *21*(2), pp. 110-118.

Kiyaga-Mulindwa, D. (1980), The" Akan" Problem. *Current Anthropology*, 21(4), pp. 503-506.

Kiyonaga, A. and Egner, T. (2012). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, 20(2), pp. 228-242.

Kline, P. (1999), *A Handbook of Psychological Testing*, 2nd edn. London: Routledge

Klingner, J., Tversky, B., Hanrahan, P., (2011), Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology*, *48*(3), pp. 323-332.

Kloss, H. (1967), 'Abstand languages' and 'Ausbau languages'. *Anthropological linguistics*, pp. 29-41.

Knoll, J, L. (2010), The "pseudocommando" mass murderer: part II, the language of revenge. *Journal of the American Academy of Psychiatry and the Law Online*, 38(2), pp. 263-272.

Konting, M., Norfaryanti, K., Man, N. (2009), Quality Assurance in Higher Education Institutions: Exist Survey among Universiti Putra Malaysia Graduating Students. *International Education Studies*. 2, pp. 297–334.

Koppensteiner, W., Lenz, A. (2020), Tracing a standard language in Austria using methodological microvariations of Verbal and Matched Guise Technique. *Linguistik online*, 102(2), pp. 47-82.

Koroutchev, R. (2017), The current refugee crisis and the challenges for Bulgaria as an entry gate for the European Union. *Journal of Liberty and International Affairs*, 3(1), pp. 9-25.

Kraidy, M, M., Khalil, J, F. (2018), *Arab Television Industries*. Bloombury, London.

Kremin, L, V., Byers-Heinlein, K. (2021), Why not both? Rethinking categorical and continuous approaches to bilingualism. *International Journal of Bilingualism*, *25*(6), pp. 1560-1575.

Kreps, G, A., Bosworth, S, L. (2007), Organizational Adaptation to Disaster. *Handbook of Disaster Research. Handbooks of Sociology and Social Research*. Springer, New York, pp. 297-315

Kruse, J., Ammon, U. (2018), The language planning and policy for the European Union and its failures. In Lian, P.C.G, Chua, C, Taylor-Leech, K., Williams, C. (eds) *Un(intended) language planning in a globalising world: Mutliple levels of players at work.* De Gruyter, Warsaw. pp. 39-56

Kudryavtsev, E, V., Ouedraogo, L, D., Unit, J, I. (2002), Implementation of multilingualism in the United Nations System. JIU/REP.2002/11.

Kuligowski, E, D., Gwynne, S. (2010), The need for behavioral theory in evacuation modeling. In *Pedestrian and evacuation dynamics.* pp. 721-732. Springer, Berlin, Heidelberg.

Kuligowski, E,D., Omori, H. (2014), Guidance document: Emergency communication strategies for buildings. *Fire Protection Research Foundation*.

Kuperberg, G, R., Jaeger, T, F. (2016), What do we mean by prediction in language comprehension?. *Language, cognition and neuroscience*, 31(1), pp. 32-59.

Kürschner, S., Gooskens, C., Van Bezooijen, R. (2008). Linguistic determinants of the intelligibility of Swedish words among Danes. *International Journal of Humanities and Arts Computing*, 2(1-2), pp. 83-100.

Kwaik, K, A., Saad, M., Chatzikyriakidis, S., Dobnik, S. (2018a). Shami: A corpus of levantine arabic dialects. In *Proceedings of the eleventh international conference on language resources and evaluation*.

Kwaik, K, A., Saad, M., Chatzikyriakidis, S., Dobnik, S. (2018b). A lexical distance study of Arabic dialects. *Procedia computer science*, 142, pp. 2-13.

Kwon, H, B. (2010), Gender difference in speech intelligibility using speech intelligibility tests and acoustic analyses. *The journal of advanced prosthodontics*, 2(3), pp. 71-76.

La Macchia, S.T., Louis, W.R. (2016). Crowd Behaviour and Collective Action. In: McKeown, S., Haji, R., Ferguson, N. (eds) *Understanding Peace and Conflict Through Social Identity Theory. Peace Psychology Book Series*. Springer, Cham.

Lachlan, K, A., Spence, P, R. (2009), Emergency communication: A framework for planning and targeting messages. *Journal of Emergency Management*, 7(3), pp. 69-72.

Laitin, D, D., Ramachandran, R. (2016), Language policy and human development. *American Political Science Review*, 110(3), pp. 457-480.

Lam, J., Tjaden, K. (2013), Intelligibility of clear speech: Effect of instruction. *Journal of Speech, Language and Hearing Research*. 56(1), pp. 1429-1440

Landick, M. (2000), French courts and language legislation. *French Cultural Studies*. 11(31) pp. 131-148

Larkin, P, J., Dierckx de Casterlé, B., Schotsmans, P. (2007), Multilingual Translation Issues in Qualitative Research: Reflections on a Metaphorical Process. *Qualitative Health Research*, 17(4), pp. 468–476.

Laverack, P, J. (2015), The Rise of Asia and the Status of the French Language in International Law. *Chinese Journal of International Law*, 14(3), pp. 567–583.

Lavie, N. (2010), Attention, Distraction, and Cognitive Control Under Load. *Current Directions in Psychological Science*, 19(3), pp. 143-148.

Leibowitz, B. (2015), The Problems with Language Policy and Planning, *Journal of Language, Identity and Education*, 14, pp. 36-49

Leonardi, M, M, V., Tamburelli, M. (2021), The cost of ignoring degrees of Abstand in defining a regional language. In Tamburelli, M. & Tosco, M. (eds). *Contested Languages: The hidden multilingualism of Europe*, 8, pp. 87-104.

Levon, E., Sharma, D., Watt, D, J., Cardoso, A., Ye, Y. (2021), Accent Bias and Perceptions of Professional Competence in England. *Journal of English Linguistics*, 49(4), pp. 355-388.

Lewis. P, M., Simons, G, F. (2010), Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Lingustique,* 55(2), pp. 103-120

Li, D. (2004), Echoes of violence: Considerations on radio and genocide in Rwanda. *Journal of Genocide Research,* 6(1), pp. 9-27.

Li, X., and Yang, Y. (2013), How long-term memory and accentuation interact during spoken language comprehension. Neuropsychologia, 51(5), pp. 967-978.

Lin, M., Martin, P, W. (2005), Decolonisation, Globalisation Language-in-Education Policy and Practice., Clevedon: Multilingual Matters.

Lindell, M, K. (2013), Disaster studies. *Current Sociology,* 61(5-6), pp. 797-825.

Lindell, M, K., and Perry, R, W. (2003), *Communicating environmental risk in multiethnic communities*. Sage publications.

Lodge, R, A. (1991), Authority, prescriptivism and the Frenchy standard language. French language studies. 1, pp. 93-111

Loos, E. (2007), Language policy in an enacted world: The organization of linguistic diversity, Language Problems & Language Planning, 31(1), pp. 37-60

Lopez, H., and Wodon, Q. (2005), The Economic Impact of Armed Conflict in Rwanda. *Journal of African Economies*, 14(4), pp. 586-602.

Lupien, S, J., Maheu, F., Tu, M., Fiocco, A., Schramek, T, E. (2007), The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and cognition*, 65(3), pp. 209-237.

MacVicar, I., (2020). What About the Camp Followers–and their Children?. *Journal for Deradicalization*, (22), pp. 319-378.

Maestri, G., Monforte, P. (2020), Who deserves compassion? The moral and emotional dilemmas of volunteering in the 'refugee crisis'. *Sociology*, 54(5), pp. 920-935.

Magnarella, P, J. (2005), The background and causes of the genocide in Rwanda. *Journal of International Criminal Justice*, 3(4), pp. 801-822.

Maheshwari, N., Rajan, K.S. (2016), A SEMANTIC MODEL TO DEFINE INDOOR SPACE IN CONTEXT OF EMERGENCY EVACUATION. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 41 pp. 313-318

Manfredi, S., Tosco, M. (2018), Arabic in contact, now and then. In Manfredi, S. and Tosco, M (eds.) *Arabic in Contact.* John Benjamins. pp. 2-17.

Manuel, J. (2014) Crisis and emergency risk communication lessons from the Elk River spill. *Environ Health Perspect.* 2014, 122(8), pp. 214-9.

Martínez, J, C., Eng, B. (2018), Stifling stateness: The Assad regime's campaign against rebel governance. *Security Dialogue*, 49(4), pp. 235-253.

Massey, S. (2021), The Bosnian genocide and the "Continuum of Denial". In *Denial: The Final Stage of Genocide*, pp. 113-130 Routledge.

Mastro, D. E. (2015), Why the media's role in issues of race and ethnicity should be in the spotlight. *Journal of Social Issues*, pp. 71, 1-16.

Mayhorn, C, B., McLaughlin, A, C. (2014), Warning the world of extreme events: A global perspective on risk communication for natural and technological disaster. *Safety science*, *61*, pp. 43-50.

McAuliffe, K. (2009), Translation at the Court of Justice of the European Communities. In *Translation issues in language and law*, pp. 99-115 Palgrave Macmillan.

McCloy, D, R., Wright, R, A., Souza, P, E. (2015), Talker versus dialect effects on speech intelligibility: A symmetrical study. *Language and Speech*, 58(3), pp. 371-386.

McGrath, J, E. (1976), Stress and behavior in organizations. In: Dunnette MD, editor. *Handbook of industrial and organizational psychology*. Chicago: McNally R pp. 1351–95.

McKenzie, R, M., Carrie, E. (2018), Implicit–explicit attitudinal discrepancy and the investigation of language attitude change in progress. *Journal of Multilingual and Multicultural Development, 39*(9), pp. 830-844.

McMenamin, J., van der Walt, C. (2018), Macro policies and meso language planning: The case of supranational policies in Europe. In Lian, P.C.G, Chua, C, Taylor-Leech, K., Williams, C. (eds) *Un(intended) language planning in a globalising world: Mutliple levels of players at work.* De Gruyter, Warsaw. pp. 57-83

Medhat-Lecocq, H. (2017), Arabe standard et variations régionales. Quelle formation pour les traducteurs arabophones? In Medhat-Lococq (eds.) *Arabe standard et variations régionales: Quelle(s) politique(s) linguistique(s)? Quelle(s) didactique(s)?. Éditions des archives contemporaines*. pp. 113-128

Miller, C. (2005), Between accomodation and resistance: Upper Egyptian migrants in Cairo. *Linguistics.* 43(5), pp. 903-956.

Miller, C. (2007), Arabic urban vernaculars: Development and change. In Miller, C. (eds.) *Arabic in the City*, pp. 15-46 Routledge, UK.

Miller, D, E. (2020), *Becoming Human Again: An Oral History of the Rwanda Genocide Against the Tutsi.* University of California Press.

Miller, N. (2013), Measuring up to speech intelligibility. *International Journal of Language & Communication Disorders,* 48(6), pp. 601-612.

Mini-ielts. (2022), *Practise tests* Available at http://mini-ielts.com (accessed 24/07/22)

Mir, S. (2019), Colonialism, postcolonialism, globalization, and Arab culture. *Arab Studies Quarterly*, 41(1), pp. 33-58.

Mirshahidi, S. (2017), I find you attractive but I don't trust you: The case of language attitudes in Iran. *Journal of Multilingual and Multicultural Development*, 38(2), pp. 146-159.

Mohanty, A.K., 2019. Language policy in education in India. In *The Routledge international handbook of language education policy in Asia*, pp. 329-340. Routledge.

Morini, M. (2013), *The pragmatic translator: An integral theory of translation* (Vol. 14). A&C Black.

Mosco, N., Atzaba-Poria, N. (2016), In search of "the Bedouin adaptive adult" socialization goals of mothers and fathers from the Bedouin society of the Negev. *Journal of Cross-Cultural Psychology*, 47(1), pp. 54-71.

Müller M, P., Hänsel, M., Fichtner, A., Hardt, F., Weber, S., Kirschbaum, C. (2009), Excellence in performance and stress reduction during two different full scale simulator training courses: a pilot study. *Resuscitation*. 80, pp. 919–24.

Mummendey, A., Klink, A., Brown, R., (2001) Nationalism and patriotism: National identification and out-group rejection. *British Journal of Social Psychology*, *40*(2), pp. 159-172.

Munro, M, J. (1993), Productions of English Vowels by Native Speakers of Arabic: Acoustic Measurements and Accentedness Ratings. *Journal of Language and Speech*. 36(1) pp. 39-66.

Munro, M, J., Derwing, T, M. (1995), Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and speech*, 38(3), pp. 289-306.

Munro, M, J., Derwing, T, M. (1999), Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, 49, pp. 285-310.

Munro, M, J., Derwing, T, M. (2011), The foundations of accent and intelligibility in pronunciation research. *Language Teaching*, *44*(3), pp. 316-327.

Murray, I, R., Arnott, J, L. (1993), Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, *93*(2), pp. 1097-1108.

Mustafawi, E. (2017), Arabic phonology. In *The Routledge handbook of Arabic linguistics*, pp. 11-31. Routledge, UK.

Nakanishi, H., Black, J., Matsuo, K. (2014), Disaster resilience in transportation: Japan earthquake and tsunami 2011. *International Journal of Disaster Resilience in the Built Environment.*

Nassenstein, N. (2019), Kinyarwanda and Kirundi: On colonial divisions, discourses of national belonging, and language boundaries. *Modern Africa: Politics, History and Society*, 7(1), pp. 11-40.

National Research Council, (1991), *Saragosa, Texas, Tornado May 22, 1987: An Evaluation of the Warning System.* National Academies Press. pp. 1-59

Nejjari, W., Gerritsen, M., Van Hout, R., Planken, B. (2019), Refinement of the matched-guise technique for the study of the effect of non-native accents compared to native accents. *Lingua*, 219, pp. 90-105.

Nelson, C. (1982), Intelligibility and non-native varieties of English. *The other tongue: English across cultures*, 15, pp. 59-73.

Nettle, D. (1996), Language diversity in West Africa: An ecological approach. *Journal of Anthropological Archaeology*, 15(4), pp. 403-438.

Ng, K, L., Hamby, D, M. (1997), Fundamentals for establishing a risk communication program. *Health Physics*, 73(3), pp. 473-482.

Ngo, T., Akagi, M., Birkholz, P. (2020), Effect of articulatory and acoustic features on the intelligibility of speech in noise: an articulatory synthesis study. *Speech Communication*, 117, pp. 13-20.

Nkejabahizi, J, C. (2007), Qu'est-ce que le Kinyagisaka? contribution à la dialectologie rwandaise. *Etudes rwandaises*, 15, pp. 22-40.

Ó Riagáin, P. (1997), *Language Policy and Social Reproduction*, Claredon Press, Wotton-under-Edge.

O'Brien, S. (2016), Training translators for crisis communication: Translators without borders as an example. *In Mediating Emergencies and Conflicts*, pp. 85-111 Palgrave Macmillan.

O'Brien, S., Federici, F., Cadwell, P., Marlowe, J., Gerber, B. (2018), Language translation during disaster: A comparative analysis of five national approaches. *International journal of disaster risk reduction*, 31, pp. 627-636.

O'Brien, S., Federici, F.M. (2020), Crisis translation: considering language needs in multilingual disaster settings, *Disaster Prevention and Management*, 29(2), pp. 129-143

Ofuji, K., Ogasawara, N. (2018), Verbal disaster warnings and perceived intelligibility, reliability, and urgency: The effects of voice gender, fundamental frequency, and speaking rate. *Acoustical Science and Technology*, 39(2), pp. 56-65.

Okada, T., Cholii, S., Karácsonyi, D., Matsumoto, M. (2021), Communities in Fukushima and Chernobyl—Enabling and Inhibiting Factors for Recovery in Nuclear Disaster Areas. In *The Demography of Disasters*, pp. 211-232 Springer.

Omar, A., Ilyas, M. (2018), The sociolinguistic significance of the attitudes towards code-switching in saudi arabia academia. *International Journal of English Linguistics*, *8*(3), pp. 79-91.

Or, I. (2017), Language Policy and Education in the Middle East and North Africa. In: McCarty T., May S. (eds) *Language Policy and Political Issues in Education. Encyclopedia of Language and Education* (3rd ed.). Springer, Cham.

Organization of African Unity (OAU) (1963), *Charter of the Organization of African Unity*, available at: https://www.refworld.org/docid/3ae6b36024.html (accessed 24/07/2022)

Owen, C., Hemmings, L., Brown, T. (2009), Lost in translation: maximizing handover effectiveness between paramedics and receiving staff in the emergency department. *Emergency Medicine Australasia*, 21(2), pp. 102-107.

Owens, J. (2015), Arabic language history and the comparative method. *International Journal of Arabic Linguistics*, *1*(1), pp. 1-27.

Pagani, M., Furlan, R., Pizzinelli, P., Crivellaro, W., Cerutti, S., Malliani, A. (1989), Spectral analysis of R-R and arterial pressure variabilities to assess sympatho-vagal interaction during mental stress in humans. *J Hypertens*. 7, pp. 14–5.

Parkinson, D, B. (1991), Searching for modern fuṣḥa: Real-life formal Arabic. *al-'Arabiyya*, pp. 31-64.

Pattanayak, D, P. (1988), Monolingual myopia and the petals of the Indian lotus: Do many languages divide or unite a nation. *Minority education: From shame to struggle*, pp. 379-89.

Peled, Y., Bonotti, M. (2019), Sound reasoning: Why accent bias matters for democratic theory. *The Journal of Politics*, 81(2), pp. 411-425.

Perlmutter, M. (1989), Intelligibility rating of L2 speech pre-and postintervention. *Perceptual and Motor Skills*, *68*(2), pp. 515-521.

Perry, R, W. (2007), What Is a Disaster?. In Anserson, W, A. Kennedy, P,J,. Kessler, E (eds.) *Handbook of Disaster Research. Handbooks of Sociology and Social Research*. Springer, New York, pp. 1-15.

Pettit, S., Beresford, A., Knight, D., Sohn, M. (2015), Humanitarian aid logistics: A new area for the public service research agenda?. In: Z. J. Radnor, et al. eds. *Public Service Operations Management: A Research Handbook*. London: Routledge, pp. 73-94.

Phillipson, R. (2006), *Language policy and linguistic imperialism. An introduction to language policy: Theory and method*, pp. 346-361

Pietersen, M, C. (2013), *The Two Largest Industrial Disasters in History with Hazardous Material.* KW Publishers.

Pool, J. (1991), The Official Language Problem. A*merican Political Science Review*, 85(02), pp. 495-514.

Powell, R. (2016), Unbalanced bilingualism, diglossia and class division in postcolonial law. *International Journal of Legal Discourse*, *1*(2), pp. 295-318.

Pronk, T., Wiers, R.W., Molenkamp, B., Murre, J., (2020), Mental chronometry in the pocket? Timing accuracy of web applications on touchscreen and keyboard devices. *Behavior Research Methods*, *52*(3), pp. 1371-1382.

Prunier, G. (2011), *Africa's World War: Congo, the Rwandan Genocide, and the Making of a Continental Catastrophe.* Oxford University Press, USA

Qafisheh, H, A. (1974), From Gulf Arabic into Modern Standard Arabic: A pilot study. *an-Nashra*, *7*(1), pp. 36-39.

Quarantelli, E, L. (1991), *Different types of disasters and planning implications.*

Quilici, A, P., Pogetti, R, S., Fontes, B., Zantut, L, F, C., Chaves, E, T., Birolini, D. (2005), Is the advanced trauma life support simulation exam more stressful for the surgeon than emergency department trauma care? *Clinics*. 60(4), pp. 287–92.

Raghubar, K, P., Barnes, M, A., Hecht, S, A. (2010), Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and individual differences*, 20(2), pp. 110-122.

Rajan, N. (2019), What do refugees want? Reading refugee lip-sewing protests through a critical lens. *International Feminist Journal of Politics*, 21(4), pp.527–543.

Ramirez, D., Engel, K, G., Tang, T, S. (2008), Language interpreter utilization in the emergency department setting: a clinical review. *Journal of health care for the poor and underserved*, 19(2), pp. 352-362.

Rawson, D. (2018), *Prelude to genocide: Arusha, Rwanda, and the failure of diplomacy*. Ohio University Press.

Ready, C. (2018), Maintaining the status quo: Diglossia and the case of Arabic language policy in Ceuta. *Language Problems and Language Planning,* 42(2), pp. 173-195.

Reyntjens, F. (2009), *The great African war: Congo and regional geopolitics*, 1996-2006. Cambridge University Press.

Rogers, C, L., DeMasi, T, M., Krause, J, C. (2010), Conversational and clear speech intelligibility of/bVd/syllables produced by native and non-native English speakers. *The Journal of the Acoustical Society of America*, *128*(1), pp. 410-423.

Ruiz, R. (1984), Orientations in language planning. *NABE journal*, 8(2), pp. 15-34.

Ruiz, R. (2010), *Reorienting language-as-resource*. In Petrovic, J, E. (eds.) *International perspectives on bilingual education: Policy, practice, and controversy.* Information Age Publishing, Charlotte, NC, pp. 155-172.

Ruiz, R. (2016), Official languages and language planning. In Hornberger, N (eds.) *Honoring Richard Ruiz and his Work on Language Planning and Bilingual Education.* Multilingual Matters Ltd, pp. 33-58.

Sadat, F., Kazemi, F., Farzindar, A. (2014), Automatic identification of arabic language varieties and dialects in social media. In *Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP),* pp. 22-27.

Saeed, A., and Sasangohar, F. (2017), Investigating the Role of Working Memory Components in Mathematical Cognition in Children: A Scoping Review. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* 61, pp. 417-421.

Saiegh-Haddad, E. (2003), Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, 24(3), pp. 431-451.

Saiegh-Haddad, E. (2004), The impact of phonemic and lexical distance on the phonological analysis of words and pseudowords in a diglossic context. *Applied Psycholinguistics*, 25(4), pp. 495-512.

Saiegh-Haddad, E. (2007), Linguistic constraints on children's ability to isolate phonemes in Arabic. *Applied Psycholinguistics,* 28(4), pp. 607-625.

Saiegh-Haddad, E., Levin, I., Hende, N., and Ziv, M. (2011), The linguistic affiliation constraint and phoneme recognition in diglossic Arabic. *Journal of Child Language*, *38*(2), pp. 297-315.

Saiegh-Haddad, E., Lina, H, A, J. (2018), Does phonological distance impact quality of phonological representations? Evidence from Arabic diglossia. *Journal of Child Language*, 45(6), pp. 1377-1399.

Samin, N. (2012), Saudi Arabia, Egypt, and the social media moment. *Arab Media & Society*, 15(1), pp. 46-65.

Samuelson, B, L., Freedman, S, W. (2010), Language policy, multilingual education, and power in Rwanda. *Language policy*, *9*(3), pp. 191-215.

Sandi, C., 2013. Stress and cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(3), pp. 245-261.

Sandman, P, M. (1989), *Hazard versus outrage in the public perception of risk. In Effective risk communication,* pp. 45-49.

Sandman, P, M. (1993), *Responding to community outrage: Strategies for effective risk communicatio*n. AIHA.

Sandman, P. (2012), Responding to Community Outrage: Strategies for Effective Risk Communication. *American Industrial Hygiene Association.* Fairfax, VA, USA.

Sankoff, G. (2016), *The social life of language*. University of Pennsylvania Press.

Santos, A., Gorter, D., Cenoz, J. (2017), Communicative anxiety in the second and third language. *International Journal of Multilingualism*, 14(1), pp. 23-37.

Scanlon, J. (2007), Unwelcome Irritant or Useful Ally? The Mass Media in Emergencies. In Anserson, W, A. Kennedy, P,J,. Kessler, E (eds.) *Handbook of Disaster Research. Handbooks of Sociology and Social Research*. Springer, New York, pp. 413-429

Schiffman, H, F. (1998), *Linguistic culture and language policy*, London; New York: Routledge.

Schoel, C., Roessel, J., Eck, J., Janssen, J., Petrovic, B., Rothe, A., Rudert, S, C., Stahlberg, D. (2013), "Attitudes Towards Languages"(AToL) Scale: A Global Instrument. *Journal of Language and Social Psychology*, 32(1), pp. 21-45.

Schulman, S.,(2019), No Exit: Outside the Camp, Daesh Incidents are on the Rise. Inside, Tempers are Seething. *The RUSI Journal*, *164*(7), pp. 54-67.

Schüppert, A., Gooskens, C. (2012), The role of extra-linguistic factors in receptive bilingualism: Evidence from Danish and Swedish pre-schoolers*. International Journal of Bilingualism*, 16(3), pp. 332-347.

Schüppert, A., Hilton, N, H., Gooskens, C. (2015), Swedish is beautiful, Danish is ugly? Investigating the link between language attitudes and spoken word recognition. *Linguistics*, 53(2), pp. 375-403.

Schüppert, A., Hilton, N, H., Gooskens, C. (2016), Why is Danish so difficult to understand for fellow Scandinavians? *Speech Communication*, 79, pp. 47-60.

Segal, U, A., Mayadas, N, S. (2005), Assessment of issues facing immigrant and refugee families. *Child welfare,* pp. 563-583.

Shaluf, I, M. (2007), Disaster types. *Disaster Prevention and Management: An International Journal.*

Shin, H., Kubota, R. (2010), Post-colonialism and Globalization in Language Education. In Spolsky, B., Hult, F., M. (eds) *The handbook of educational linguistics*, pp. 206-220.

Shouby, E. (1951), The influence of the Arabic language on the psychology of the Arabs. *The Middle East Journal*, pp. 284-302.

Skeldon, R. (2009), Migration and Development: Contested Consequences. In Kremer, M., van Lieshout, P,. Went, R. (eds). *Doing Good Or Doing Better: Development Policies in a Globalizing World*. Amsterdam University Press, pp. 321-340.

Skutnabb-Kangas, T. (2009), Language Policy and Linguistic Human Rights. In: T. Ricento, ed. *An Introduction to Language Policy: Theory and Method*. London: Johnn Wiley & Sons, pp. 273-290.

Skutnabb-Kangas, T., 2006. Language policy and linguistic human rights. *An introduction to language policy: Theory and method*, pp. 273-291.

Smith, A, D. (2013), Nationalism: Theory, ideology, history. John Wiley & Sons.

Smith, L, E., and Nelson, C, L. (2019), World Englishes and issues of intelligibility. In Nelson, C.L., Proshina, Z.G. & Davis, D.R (eds.), *The handbook of world Englishes*, pp. 430-446.

Smith, W, D., Chung, Y, H., Berguer, R. (2000), A virtual instrument ergonomics workstation for measuring the mental workload of performing video-endoscopic surgery. *Stud Health Technol Inform*, 70, pp. 309–15.

Snow, D. (2013), Revisiting Ferguson's defining cases of diglossia. *Journal of Multilingual and Multicultural Development*, 34(1), pp. 61-76.

Sokolovska, Z. (2017), Language in "the United Nations of Europe": debating a post war language policy for Europe. *Lang Policy*. 16, pp. 461-480

Sorensen, J.H., Sorensen, B.V., 2007. Community processes: Warning and evacuation. In *Handbook of disaster research*, pp. 183-199.

Sourander, A. (2003), Refugee families during asylum seeking. *Nordic Journal of Psychiatry*, 57(3), pp.203-207.

Spolsky, B. (2011), Language academies and other language management agencies. *Language Policy*, 10(4), pp. 285-287.

Staal, M, A. (2004), *Stress, cognition, and human performance: A literature review and conceptual framework.*

Stadlbauer, S. (2010), Language Ideologies in the Arabic Diglossia of Egypt. *Colorado Research in Linguistics*, *22*.

Starcke, K., Brand, M. (2012), Decision making under stress: a selective review. *Neuroscience & Biobehavioral Reviews*, 36(4), pp. 1228-1248.

Staub, E. (2009), The Origins of Genocide and Mass Killings: Core Concepts. In: S. Totten & P. R. Bartrop, eds. *The Genocide Studies Reader*. London: Routledge, pp. 97-108.

Straus, S. (2006), *The Order of Genocide: Race, Power, and War in Rwanda.* Cornell University Press.

Suied, C., Susini, P., and McAdams, S. (2008), Evaluating warning sound urgency with reaction times. *Journal of experimental psychology: applied*, *14*(3), pp. 201-212.

Sutton, J., Spiro, E, S., Johnson, B., Fitzhugh, S., Gibson, B., Butts, C, T. (2014), Warning tweets: Serial transmission of messages during the warning phase of a disaster event. Information, *Communication & Society*, 17(6), pp. 765-787.

Sweller, J., Ayres, P., Kalyuga, S. (2011), *CLT theory*. New York: Springer

Taber, K, S. (2018), The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Res Sci Educ* 48, pp. 1273–1296.

Taibi, M., Ozolins, U. (2016), *Community translation*. Bloomsbury Academic, London.

Tajfel, H. (1979), Individuals and groups in social psychology. *British Journal of social and clinical psychology*, 18(2), pp. 183-190.

Tajfel, H. (2001), Social stereotypes and social groups. *Intergroup relations: Essential readings*, pp. 132-145.

Tajfel, H., Turner, J.C. (2004), The social identity theory of intergroup behavior. In *Political psychology*, pp. 276-293.

Tajfel, H., Turner, J.C., Austin, W.G. and Worchel, S., (1979). An integrative theory of intergroup conflict In Hatch, M., Schultz., M (eds) *Organizational identity: A reader*, pp. 55-65.

Tajfel, H.,Turner, J, C. (2004), The social identity theory of intergroup behavior. In *Political psychology*, pp. 276-293.

Tamburelli, M. (2014), Uncovering the 'hidden' multilingualism of Europe: an Italian case study. *Journal of Multilingual and Multicultural Development*, 35(3), pp. 252-270.

Tamburelli, M. (2021), Taking taxonomy seriously in linguistics: Intelligibility as a criterion of demarcation between languages and dialects. *Lingua*, 256, pp. 103068.

Tan, K, H., Perudin, A. (2019), The "Geopolitical" factor in the Syrian Civil War: A corpus-based thematic analysis. *Sage Open*, 9(2), pp. 1-15.

Tang, C. (2009), *Mutual intelligibility of Chinese dialects : an experimental approach*. Utrecht: Leiden University.

Tang, C., Van Heuven, V, J. (2007), Mutual intelligibility and similarity of Chinese dialects: Predicting judgments from objective measures. *Linguistics in the Netherlands*, Volume 12, pp. 223-234.

Tang, C., Van Heuven, V, J. (2009), Mutual intelligibility of Chinese dialects experimentally tested. *Lingua*, 119(5), pp. 709-732.

Tang, C., Van Heuven, V, J. (2015), Predicting mutual intelligibility of Chinese dialects from multiple objective linguistic distance measures. *Linguistics*, 53(2), pp. 285-312.

Tannenbaum-Baruchi, C., Feder-Bubis, P., Aharonson-Daniel, L. (2017), Dialogue in Emergencies - Interpreters of Sign Language in Israel. *Prehospital and Disaster Medicine*, *32*(S1), pp. S181-S182.

Tanner, L., Obrecht, A. (2015), 'Words of Relief: Translators without Borders' local language

Teigen, K.H., 1994. Yerkes-Dodson: A law for all seasons. *Theory & Psychology*, *4*(4), pp. 525-547.

Theodoropoulou, I. (2015), Sociolinguistic anatomy of mobility: Evidence from Qatar. *Language & Communication*, *40*, pp. 52-66.

Theodoropoulou, I., Tyler, J. (2014), Perceptual dialectology of the Arab world: a principal analysis. *Al-ʿArabiyya*, pp. 21-39.

Thorndike, E, L., Lorge, I. (1952), *The Teacher's Word Book of 30,000 Words.* Teachers College, Columbia University

Todd, J., and Marois, R. (2004), Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, 428(6984), pp. 751-754.

Tollefson, J, W. (2006), Critical Theory in Language Policy. In: *An Introduction to Language Policy: Theory and Method.* London: John Wiley & Sons.

Tomaka, J., Blascovich, J., Kelsey, R, M., Leitten, C, L. (1993), Subjective, physiological, and behavioral effects of threat and challenge appraisal. *J Pers Soc Psychol.* 65, pp. 248–60.

Tomasi, L. (2017), The European Commission's language policy: challenges and priorities. *European Journal of Language Policy*, *9*(1), pp. 121-131.

Tonkin, H. (2011), Language and The United Nations: A Preliminary Review. *Léger*, 1(1), pp. 2-5.

Tosco, M. (2008), Introduction: Ausbau is everywhere!. *International Journal of the Sociology of Language*, 161, pp. 1-16.

Tosco, M. (2011), *Between endangerment and Ausbau.*

Tosco, M. (2017), *On counting languages, diversity-wise.*

Tosco, M., Tamburelli, M. (2021), Contested Languages. *Contested Languages*, pp. 1-277.

Towler, M. (2018), Language Mapping Pilot Project: Arabic Language Provision in the UK.

Translators without Borders. (2017a), Putting Language on the Map in the European refugee response. *Mixed Migration Platform*.

Translators without Borders. (2017b), Lost for Words: Improving information access for refugees and migrants in Greece. *Mixed Migration Platform.*

Translators without Borders. (2017c), *A Study on the Multitude of Languages and Comprehension of Material Provided to Refugees and Migrants in Greece. Language & Comprehension Barriers in Greece's Migration Crisis.*

Translators without Borders. (2017d), A Study on the Impact of Language Barriers on Refugee and Migrant Children in Greece. *BRIDGING THE GAP.*

Translators without Borders. (2017e), *The field guide to humanitarian interpreting and cultural mediation.*

Trentman, E., Shiri, S. (2020), The mutual intelligibility of Arabic dialects: Implications for the language classroom. *Critical Multilingualism Studies*, *8*(1), pp. 104-134.

Trudgill, P. (1992), Ausbau sociolinguistics and the perception of language status in contemporary Europe. *International Journal of Applied Linguistics*, 2(2), pp. 167-177.

Trudgill, P. (2009), Contact, isolation, and complexity in Arabic. In Muraoka, T., Versteegh, C. (eds) *Arabic dialectology* , pp. 171-186.

Tubi, A., Feitelson, E. (2016), Drought and cooperation in a conflict prone area: Bedouin herders and Jewish farmers in Israel's northern Negev, 1957–1963. *Political Geography*, 51, pp. 30-42.

Turner, S., 2016. What is a refugee camp? Explorations of the limits and effects of the camp. *Journal of Refugee Studies*, *29*(2), pp. 139-148.

Uchanski, R. (2005), Clear Speech. In Pisoni, D,B., Remez, R, E. (eds) *The Handbook of Speech Perception*. Blackwell Publishing, London, pp. 207-235

UNESCO. (2003), Language Vitality and Endangerment. *UNESCO Ad Hoc Expert Group on Endangered Languages*.

United Nations (2022), *Member States*. Available at: https://www.un.org/en/about-us/member-states (Accessed 24/07/22)

United Nations Committee on Economic, Social and Cultural Rights (CESCR). (1999), 'General comment no. 13: The right to education (article 13)'. United Nations Committee on Economic and Cultural Rights at the Twenty-first Session, E/C.12/1999/10.

United Nations General Assembly. (1946), Rules of Procedure Concerning Languages. UNGA 7; A/RES/2 (I)

United Nations General Assembly. (1948), *Proposal for the adoption of Spanish as one of the working languages of the General Assembly. Resolution A/RES/247(III).*

United Nations General Assembly. (1968), *Inclusion of Russian among the working languages of the General Assembly (amendment to rule 51 of the rules of procedure of the General Assembly) and question of including Russian and Spanish among the working languages of the Security Council.* Resolution A/RES/2479(XXIII)

United Nations General Assembly. (1973), *Inclusion of Arabic among the official and the working languages of the General Assembly and its Main Committees.* A/RES/3190(XXVIII)

United Nations General Assembly. (1973), *Inclusion of Chinese among the working languages of the General Assembly and the Security Council. A/RES/3189(XXVIII)*

United Nations, Charter of the United Nations. (1945), 1 UNTS XVI

United Nations, *Charter of the United Nations.* (1945), 1UNTS XVI, available at: https://www.un.org/en/charter-united-nations/index.html [accessed 24/07/2022]

Van der Borgh, C. (2009), Post-war Peace-building: What Role for International Organizations? In Kremer, M., van Lieshout, P,. Went, R. (eds). *Doing Good Or Doing Better: Development Policies in a Globalizing World*. Amsterdam University Press, pp. 303-320

Van der Borgh, C., (2009) Post-war peace-building: what role for international organizations?. In Kremer, M., van Lieshout, P., Went, R., *Doing Good or Doing Better*, pp. 303-320

van der Velden, P, G., Yzermans, C, J., Grievink, L. (2012), *Enschede Fireworks disaster*. Cambridge University Press.

Van Heuven, V, J. (2008), Making sense of strange sounds:(Mutual) intelligibility of related language varieties. A review. *International journal of humanities and arts computing*, 2(1-2), pp. 39-62.

Van Heuven, V, J., Gooskens, C, S., Bezooijen, R, V., Navracsics J, B, S. (2015), Introducing Micrela: Predicting mutual intelligibility between closely related languages in Europe. *Studies in Psycholinguistics*, pp. 127-145.

Van Merriënboer, J, J., Kirschner, P, A., Kester, L. (2003), Taking the load off a learner's mind: Instructional design for complex learning. *Educational psychologist*, 38(1), pp. 5-13.

Versteegh, C, H, M. (1993), *Arabic grammar and Qur'ānic exegesis in early Islam* (Vol. 19). Brill.

Versteegh, K. (2013), *Landmarks in linguistic thought volume III: The Arabic linguistic tradition*. Routledge.

Versteegh, K. (2014), *Arabic language*. Edinburgh University Press.

Vij, R. (2022), Types of Disasters. In *Management of Animals in Disasters*, pp. 3-14 Springer.

Villada, C., Hidalgo, V., Almela, M., Salvador, A. (2016), Individual differences in the psychobiological response to psychosocial stress (Trier Social Stress Test): the relevance of trait anxiety and coping styles. *Stress and Health*, 32(2), pp. 90-99.

Walters, K. (1996), Diglossia, linguistic variation, and language change in Arabic. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES,* 4, pp. 157-200.

Walther, O, J., Pedersen, P, S. (2020), Rebel fragmentation in Syria's civil war. *Small Wars & Insurgencies,* 31(3), pp. 445-474.

Warheit, G, J., Zimmerman, R, S., Khoury, E, L., Vega, W, A., Gil, A, G. (1996), Disaster Related Stresses, Depressive Signs and Symptoms, and Suicidal Ideation Among a Multi-Racial/Ethnic Sample of Adolescents: A Longitudinal Analysis. *Journal of Child Psychology and Psychiatry*, 37: pp. 435-444

Watson, J, C. (2007), *The phonology and morphology of Arabic*. OUP Oxford.

Weiss, M. (2015), *The comparative method. In The Routledge handbook of historical linguistics* pp. 127-145. Routledge.

Weiss, T, G. (2010), How United Nations ideas change history. *Review of International Studies*, *36*(S1), pp. 3-23.

Wetzel, C, M., Black, S, A., Hanna, G, B., Athanasiou, T., Kneebone, R, L., Nestel, D., Wolfe, J, H., Woloshynowych, M. (2010), The effects of stress and coping on surgical performance during simulations. *Annals of surgery*, 251(1), pp. 171-176.

Wiley, T, G., García, O. (2016), Language Policy and Planning in Language Education: Legacies, Consequences, and Possibilities. *Modern Language Journal,* 100(S1), pp. 48–63.

Williams, B., Cassar, C., Siggers, G., Taylor, S. (2016), Medical and social issues of child refugees in Europe. *Archives of Disease in Childhood*, 101(9), pp. 839-842.

Williams, C, H. (2000), *Language revitalization: policy and planning in Wales,* University of Wales Press.

Williams, C, H. (2014), The Lightening Veil: Language Revitalization in Wales. *Review of Research in Education*, 38(1), pp. 242–272.

Williams, E. (2011), Language policy, politics and development in Africa. *Dreams and realities: Developing countries and the English language*, pp. 39-56.

Wogalter, M, S., Conzola, V, C., Smith-Jackson, T, L. (2002), based guidelines for warning design and evaluation. *Applied ergonomics*, *33*(3), pp. 219-230.

World Bank (2019). *Population [Total]*. Available at https://data.worldbank.org/ (Accessed: 24/07/2022)

Wright, S. (2016), Language Policy and Planning: FROM NATIONALISM TO GLOBALISATION. Palgrave Macmillian. 2nd Edition.

Yorkston, K, M., Strand, E, A., Kennedy, M, R. (1996), Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology*, 5(1), pp. 55-66.

Youssef, I. (2019), The phonology and micro-typology of Arabic R. *Glossa: a journal of general linguistics* 4(1): 131, pp. 1–36.

Yusof, Z, M., Hussain, R., Ahmed, M. (2013), Malay Speech Intelligibility Test (MSIT) for Deaf Malaysian Children. *International Journal of Integrated Engineering*, 5(3).

Zaidan, O, F., Callison-Burch, C. (2014), Arabic dialect identification. *Computational Linguistics*, 40(1), pp. 171-202.

Zaidan, O., Callison-Burch, C. (2011), The arabic online commentary dataset: an annotated dataset of informal arabic with high dialectal content. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: *Human Language Technologies*, pp. 37-41.

Zaidan, O., Callison-Burch, C. (2012), Machine translation of Arabic dialects. *Proceedings of the 2012 conference of the north American chapter of the association for computational linguistics: Human language technologies*, pp. 49-59.

Zbib, R., Malchiodi, E., Devlin, J., Stallard, D., Matsoukas, S., Schwartz, R., Makhoul, J., Zaidan, O., Callison-Burch, C. (2012), Machine translation of Arabic dialects. In *Proceedings of the 2012 conference of the north american chapter of the association for computational linguistics: Human language technologies*, pp. 49-59.

Żuber, M., Moussa, S, S. (2018), Arab spring as a background of civil war in Syria. In *International Conference Knowledge-Based Organization* Vol. 24, pp. 245-251.