CHAPTER I

INTRODUCTION

1.1 Background

Language can be described as a system of arbitrary, agreed-upon symbols used by a community for the purposes of communication. But it is also much more than this. Language is a symbol of group distinctiveness, and a vehicle of tradition and culture (Edwards, 2009, p. 54). In fact, language is one of the most powerful instruments for the construction and expression of identity, both on an individual and a group level. As such, languages can be a source of strong feelings, and so have been both the ammunition and the casualties of political battles around the world (B. Anderson, 2006; Tabouret-Keller, 1998). On the other hand, when people are ashamed of their identity as a member of a non-prestigious minority group, this can lead to minority languages growing weak and dying out as their speakers shift to majority or prestige language use (Romaine, 2010). The study of language and identity provides a fascinating insight into culture and humanity. Sulawesi, the third most linguistically diverse island in the world (Mead, 2013, p. 3), is a complex tapestry of interconnected languages and dialects with rich cultural history, and yet broad variation in attitudes towards language identity, thus making it an ideal location for a study of language and identity.

One area of South Sulawesi that is of significant interest in this regard is Luwu, the northernmost region of South Sulawesi province. Zainal records that Couvreur, Governor of Celebes in 1929, claimed Luwu was the most powerful Sulawesi kingdom between the tenth and fourteenth centuries, and that in fact the genealogies of the kings and nobility of South Sulawesi show that Luwu was founded before the kingdoms of Bugis, Mandar and Makassar (1983, p. 462). However, by the late 1800s, Luwu no longer held the political and economic power it once had (Errington, 2014), and ongoing power struggles and extensive migration to the area led to confusion over a Luwu ethnic and linguistic identity (Both, 2005, p. 497). The original language of the Luwu kingdom and its several [illegible] have been ascribed various names by different linguists (e.g. Adriani, [illegible]rr et al., 1979; Esser, 1938; Grimes & Grimes, 1987; Kruyt, 1920; Salzner, K. Sirk, 2007; Sneddon, 1981; Vail, 1991a; Veen, 1929), and to this day

there is still no clear consensus on the name of the language, with various researchers describing it as Bahasa Tae' (Garing, 2011; Ibrahim, 2013; Idawati et al., 2016a; Sahraini, 1998) or Bahasa Taeq (Garing, 2012, 2014), while Laskowske and Mead observe that locals tend to call it Bahasa Luwu' (2004).

To date, relatively little linguistic research has focused on the Luwu' language. While there are numerous works published concerning other South Sulawesi languages, e.g. Bugis (Abas & Grimes, 1995; Achmad, 2012; Friberg & Friberg, 1988; Junus & Junus, 2007; Kaseng, 1975; D. C. Laskowske, 2011; Mahmud, 2008a, 2008b, 2010; Mokhtar, 2000; Nuraidar, 2012, 2014; Sikki, 1991; J. X. Sirk, 1983; Valls, 2014); Toraja (Biring, 1981; J. Jerniati, 2010; Machmoed, 2012; Salombe, 1978, 1979, 1982; Sande, 1976, 1984; Sande & Stokhof, 1977; Sikki, 1986; Yatim & Machmoed, 2007); and Makassar (Arisnawati, 2012; Arsyad et al., 1987; Basang et al., 1975; Daeng & Syamsuddin, 2005; Hasyim, 2008; Jukes, 2013b, 2013a, 2015b, 2015a; Kaseng, 1978; Manyambeang, 1979; Manyambeang et al., 1996; Marshall, 2018; Tabain & Jukes, 2016), Luwu' remains somewhat of a mystery. In 1988, Ian Vail carried out lexicostatistic research throughout Kabupaten Luwu (which at that time included the current districts of Luwu, Palopo, North Luwu and East Luwu), but he admitted that his conclusions were not definitive, and highlighted general tendencies rather than specific details (1991a, p. 79). Since then, there have been some small-scale studies on language maintenance and shift among Palopo youth, and among Tae' speakers living in Makassar (Djamareng & Jufriadi, 2016; Rusli, 2016), and on various grammatical aspects of Tae' or specific Tae' dialects (e.g. Garing, 2011; Ibrahim, 2013; Idawati et al., 2016a; Sahraini, 1998; Salija et al., 2017).

There remains then, a wide open door for research into the area of language and identity in the Luwu region. I propose to research this issue from several facets to provide a comprehensive picture of the status of the Luwu' language and its relation to a Luwu identity, including dialect intelligibility testing, sociolinguistic surveys, and some unique aspects of Luwu' morphology.

A further aspect related to language and identity is the effect second language learning has in shaping the learner's identity. Identity is "socially constituted, a reflexive, dynamic product of the social, historical and political of an individual's lived experiences" (Hall, 2012, p. 31). In the process of ng language and identity in Luwu, I am also interested in recording the arning the language has on my own identity. Rarely do researchers

document their own ability in the language under study and the effect this ability or lack thereof has on the quality of their research (Borchgrevink, 2003; Gibb & Danero Iglesias, 2017). Therefore I aim to address this shortfall in a small way by adding an academic reflection on my own experiences to the existing literature (e.g. Moore, 2009; Rodgers, 2012; Tremlett, 2009).

1.2 Research Focus

In the following subsections I explain the focus of each aspect of my research.

1.2.1 Sociolinguistic surveys

Identification with one's ethnic group and its language has a major impact on individual language maintenance, and on a group level, on language vitality (Tabouret-Keller, 1998). As Djamareng and Jufriadi also note in their research on Tae' language maintenance in Makassar, positive attitudes towards a minority language by its speakers influence the language's status (2016, p. 82). Thus sociolinguistic surveys on language attitudes, domains of use, and multilingualism can be used to assess ethnolinguistic vitality and language status. In the course of his research in the 1980s, Vail administered an oral questionnaire throughout the Luwu region on language use and literacy, and also mapped demographic patterns throughout the area, including language contact and patterns of use (1991a). However, in the ensuing thirty years, which have seen tremendous growth in infrastructure, mobility and technology, there has been no further published sociolinguistic research on Luwu' language vitality in Luwu outside of Palopo, and this research seeks to meet that need. It asks:

- What are the attitudes of Luwu people to their language?
- What patterns of use are displayed in different generations, regions and domains?
- What connection do these patterns and attitudes have to a Luwu identity?
- In which domains of use is Luwu identity the strongest?

alect intelligibility testing

uring the 1970s and 80s, a number of linguists carried out broad-scale apping throughout South Sulawesi using lexicostatistical methods. Their

research resulted in the current nomenclature of South Sulawesi languages. However in the course of their research, they acknowledged the difficulties they encountered using lexicostatistics alone. Chaining patterns among the dialects and languages of South Sulawesi make it difficult to classify them hierarchically (Grimes & Grimes, 1987, p. 13). Simons states that in such instances, lexicostatistic data is not sufficient (1977, p. 116). Indeed, upon the conclusion of his lexicostatistic dialect mapping in Luwu, Vail wrote, "[t]here remain a number of foci for future study, not the least of which is to sort out the intelligibility of these dialects to one another" (1991a, p. 93). This 30 year old call for further research has not yet been met, but it is a crucial puzzle piece in the picture of the Luwu ethnic and linguistic identity.

As Bucholtz and Hall point out, there are two aspects to identity categorisation, namely the objectively describable speech community, and the subjective socio-cultural personal attitudes towards identity of the group members themselves (2004, p. 370). Thus the first step in any identity work here must be a clear understanding of the linguistic situation, which can then be correlated with the aforementioned sociolinguistic survey results for a more nuanced view of a Luwu identity. This research asks therefore:

- How mutually intelligible are the dialects of Luwu'?
- What relation does this have to a broader Luwu identity?

1.2.3 Morphological analysis

In addition to the big-picture aspects of objective dialect boundaries correlated with subjective attitudes towards group identity, I will home in on two examples of Luwu' morphology that speakers regularly employ to convey fine shades of meaning and emphasis in the stories they tell. The first feature is the infix *-um-*, which is also widespread throughout the Austronesian language family, and yet varies markedly in function (Blust & Trussel, 2020). The second feature is emphasis for intensity in stative verbs, which is shared by several surrounding and closely related languages, yet not present at all in other neighbouring languages. These two morphological features of Luwu' showcase the complex relationship between Luwu' and the broader South Sulawesi and Austronesian language family.

ɜ:

'hat are the function and purpose of these two morphological features?

- Do these morphological features vary between different dialects and group identities?

1.2.4 Identity reflection

Finally, regarding identity development through second language acquisition, most research to date focuses on the language classroom (McKay & Wong, 1996; Norton, 2000; Norton & Toohey, 2011). Most researchers also focus on those who are learning a language of greater global currency than their mother tongue, either in an effort to gain access to greater economic, social or political resources, or under pressure to conform in order to survive (Johnston, 1999; Norton, 2000). However, my interest lies in another direction. I am an English mother-tongue speaker, learning a minority language with no classroom input.

Furthermore, a key aspect in the discussion of language and identity is the notion of sameness and difference, or 'adequation' and 'distinction', to use Bucholtz and Hall's terminology (2004, p. 383). Group identity is not a fixed or intrinsic state, rather it is a process of "inventing similarity by downplaying difference" (Bucholtz & Hall, 2004, p. 371). Temple and Edwards use the concept of borders to discuss identity (2002, p. 8), noting that there are no clear lines between "us" and "them", rather there are many ways in which individuals and groups differ, and different aspects are more or less salient at different times. Individuals are constantly negotiating and crossing these borders as they interact and establish relations with others. Therefore, as my identity as the foreign researcher in relation to the local community continually changes due to my growing language ability, the following questions arise:

- What impact does my language ability have on data collection?
- How relevant is linguistic ability in terms of group membership in the eyes of the local community?

1.3. Research Aims

The sociolinguistic surveys conducted will provide crucial up-to-date information for government and community language planning. As Lewis and Simons explain, for a community to make informed decisions about their language nce and development goals, outside expertise providing scientifically language vitality information is an important prerequisite (2017, p. 14). nrough dialect intelligibility testing I aim to determine the degree of ity between two speech varieties in South Luwu.

I will also seek to analyse two discrete morphological features in Luwu' to highlight the complex relationship between Luwu' and the surrounding languages.

And in presenting an academic reflection on my personal memoirs, I hope to provide valuable insights to the field of language learning and identity.

1.4 Expected Outcomes

The multi-faceted, interdisciplinary nature of this research will provide many practical outcomes both for the Luwu region and beyond:

Firstly, data from the sociolinguistic surveys will assist in government language planning and policy decisions regarding the promotion of Luwu' language maintenance in the home and community.

Secondly, the information on dialect intelligibility across the different regions of Luwu produced by this research will assist in developing a standard modern orthography, and the development of literacy materials that will be widely accessible to all Luwu' language speakers. Furthermore, the implementation of Yoder's (2014) suggested modifications of the traditional RTT methodology (see section 3.6.2 below) will provide a useful cross-reference for researchers in other parts of the world in the ongoing effort to develop more robust and accurate dialect intelligibility testing methods.

Thirdly, it is anticipated that the details unearthed in the process of morphological analysis will form a basis for further research in other South Sulawesi languages, and will also enhance our understanding of some of the shared features of Austronesian languages right across the region.

Finally, it is hoped that my personal language acquisition record will lead by example to greater transparency among researchers in reporting their language abilities and publishing their own language learning memoirs, and add to the corpus of such texts to enable comparative analysis (Gibb & Danero Iglesias, 2017, p. 145). In addition, it will provide a timely response to Norton and Toohey's call for a focus on identity and SLA processes "in postcolonial and indigenous sites, where multilingualism is ubiquitous and language acquisition processes can be quite different from language learning experiences in the West" (2011, p. 436).

In all, the Luwu' language region is a promising area for much fruitful on language and identity, uncovering fascinating insights into culture and .

1.5 Research Scope

This research focuses specifically on the Luwu' spoken in the Luwu Regency, Ian Vail's South Luwu dialect (1991a, p. 79). The Rongkong, North-East Luwu, and Bua dialects spoken in the northern part of the Luwu Regency, North Luwu Regency, and East Luwu Regency, are excluded from this study.

1.6 Research novelty

Each subsection of this research is novel in its own right, as is the combination into the whole as a way of considering the topic of language and identity. Firstly, language use and attitudes surveys of this kind have not been done in this area in at least 30 years. Secondly, dialect intelligibility testing using recorded texts has not been done in this language before. Thirdly, the morphological features discussed, while present to varying degrees in some neighbouring and related languages, have never been methodically or accurately explained in Luwu'. And finally, while a small number of other researchers have recorded and written about their own language ability and identity in relation to their research, my own observations are unique in that they relate to a new subject: myself.

CHAPTER II
LITERATURE REVIEW

2.1 Review of Theories and Concepts

The relationship between language and identity has long been acknowledged, with records as far back as the tenth century BC in the famous *shibboleth* test, where the people of Gilead used soldiers' pronunciation of this one word to determine their ethnic group membership – and immediately killed all enemies thus identified (Tabouret-Keller, 1998). However, only far more recently has language and identity become a topic of study in its own right, as "up to the mid-1990s, there had been little if any research in the [applied linguistics] field which either cited identity or employed it as a central construct" (Block, 2013, p. 13). And yet, as Joseph (2004) declares,

> ...any study of language needs to take consideration of identity if it is to be full and rich and meaningful, because identity is itself at the very heart of what language is about, how it operates, ... how it is learned and how it is used, every day, by every user, every time it is used. (p. 224)

During the past couple of decades, the focus on language and identity has permeated many areas of the broader applied linguistics field – for example there has been work on identity and sociolinguistics (e.g. Edwards, 2009; Joseph, 2004), identity and multilingualism (e.g. Edwards, 2013; Pavlenko et al., 2001), and identity and language learning (e.g. McKay & Wong, 1996; Norton, 2000), to mention just a few areas of relevance to my own research. Despite this, there is much controversy both amongst linguists, and social scientists more broadly, over the very definition of identity itself (Block, 2013). In essence, identity is the condition of being oneself and not another, having a certain sameness and continuity over time ('Identity', 1991). However beyond this basic tenet, some see identity "as something constructed, fluid, multiple, impermanent and fragmentary" (Bendle, 2002, p. 2), with individuals being the shapers of their own realities, and identities in constant flux. Freeman (2001) even goes so far as to say that

> ...there *is* no identity to the self, no condition of permanent elfsameness. .... Or, to put the matter in more positive terms, there as emerged the recognition that personal identity is *changeable* – cross time, across space, and, more generally, across the various scursive contexts within which identity is negotiated and, on some vel, produced. (p. 295)

In contrast to this post-structuralist view, an essentialist view sees identity as a fixed and permanent, "inevitable and natural" position in society (Bucholtz & Hall, 2004, p. 374), giving no agency to the individual to develop and change. While most scholars today reject such an extreme position, critical realists hold that it cannot be denied that to an extent "we are shaped and affected by social structures" (Lopez & Potter, 2001, p. 15), as we work within, manipulate or react against the social norms around us in the expression and enactment of our Selves, and this is the perspective of identity that my research will take – and I trust, support. With such a diverse range of views, it is no wonder Bucholtz and Hall suggest that the concept of identity needs to be more fully theorised in order to strengthen further research in the field (2004, p. 387).

At this point it would be helpful to distinguish between individual or personal identity, and collective or group identity, both of which are important concepts in my research. Individual identity speaks of the concept of the personality, "the fact that a person is oneself and not someone else. It signifies a continuity ... that constitutes an unbroken thread running through the long and varied tapestry of one's life" (Edwards, 2009, p. 19). Collective identity, on the other hand, refers to a sense of unity with certain other individuals based on specific shared characteristics. This similarity within the group is emphasised as being distinct from others outside the group who do not share these same traits. As Bucholtz and Hall assert, "The perception of shared identity often requires as its foil a sense of alterity, of an Other who can be positioned against those socially constituted as the same." (Bucholtz & Hall, 2004, p. 371). Thus both individual and collective identities are enacted and constructed within a paradigm of the Self and the Other. Of course this is not to say that either individual or group identity is simple and straightforward. Individuals are multifaceted and part of many groups simultaneously – in fact, may even be part of multiple ethnic groups as a result of mixed parentage – which is why many scholars refer to identities as plural and multiplex (Hall, 2012; Lemke, 2008).

Along with the very definition of identity as mentioned above, the relationship between identity and language is also contested. Some scholars hat a specific language is intrinsically tied to a specific cultural identity, erring to the unique cultural knowledge encoded in the local language. In ge declares that "a people who lose their language, and the view of the expressed by that language, can no longer survive as a people, but only

as rootless individuals" (Bunge, 1992, p. 380; see also Skutnabb-Kangas & Dunbar, 2010). Others, such as May, hold that

> ...the languages we speak are not ineluctably linked to our (ethnic) identity. .... language does not define us, and may not be an important feature, or indeed even a necessary one, in the construction of our identities, whether at the individual or collective levels. (2013, p. 9)

Personally I adopt a stance between these two camps. Languages and ethnic groups are both constantly changing slowly over time – "culture needs to be viewed as fluid and evolving and discourse as open rather than closed" (Evans, 2015, p. 6). Ethnic groups separate and merge, languages are adapted to suit changing needs, and dialectal differences grow stronger with distance, or weaker with contact (see Peters & Fee, 1989 regarding these effects in Britain and its former colonies). Essentialist views such as those promoted by Bunge ignore these facts. Furthermore, regarding the view held by May, for some monolinguals who have little to no contact with other language communities, it is possible that language is not a conscious part of their identity – however, it is largely through language that they express their identity, whether consciously or not (Bailey, 2007, p. 341). Some individual members of ethnic groups who abandon their heritage language in favour of a more prestigious national language may not feel a strong connection between their language and identity (Edwards, 2010; Khelmani-David, 2009; Pandharipande, 1992), but for others, language is a defining feature of their group identity, and thus something for which they are willing to die or kill (Tabouret-Keller, 1998). Thus, the languages we speak are tied tightly to our identities, but their salience to us as their speakers can vary over time and between individuals and groups.

When we focus on the level of dialects in our consideration of language and identity, further complexities come to light. In the simplest terms, a dialect is a language variety that differs from other varieties of that language in systematic ways in terms of vocabulary, grammar and pronunciation, but is nevertheless mutually intelligible with other varieties of the language (Fromkin et al., 2009). However, Wolff gives several examples from his research in Nigeria of discrepancies between linguistic similarity and both reported and tested

ity, with factors such as economic power, inter-ethnic relationships, and stance also playing a significant role in dialect differentiation (1959), that identity can have a considerable impact on intelligibility. To cite one

example, Wolff writes about the speakers of the Isoko dialect of the Urhobo language. He says,

> ...until recently there was general agreement that mutual intelligibility was relatively high among all Urhobo dialects. Lately, however, speakers of Isoko have been claiming that their language is different from the rest of Urhobo, and that intelligibility between Urhobo and Isoko is not sufficient for normal linguistic communication. This claim has coincided with Isoko demands for greater political autonomy and ethnic self-sufficiency. (1959, p. 37)

Thus, the issue of identity is a key factor in studies of dialect intelligibility.

Having established the key theories and issues being debated in the broader sphere in which my research fits, I will now focus on the literature related to the specific elements of my own research in South Sulawesi.

2.2 Review of Relevant Works

Indonesia is one of the most ethnically and linguistically diverse nations in the world, thus providing great scope for studies of language and identity. Researchers have addressed the topic from various angles, including Lowenberg's discussion of the Indonesian language as a powerful force in developing a sense of national identity (1992), Kuipers' monograph on subtle changes in certain aspects of language use in Sumba and the implications for identity (1998), and Goebel's study on the co-construction of identity in multi-ethnic neighbourhoods in Java (2010). Linguistic research in Sulawesi has traditionally focused on aspects of grammar (e.g. Jukes, 2013a, 2013b, 2015a, 2015b; Quick, 2007), politeness (e.g. Mahmud, 2008a, 2010, 2013), and language and dialect mapping (e.g. Grimes & Grimes, 1987; Machmoed, 2012; Mead, 1999, 2013). As for the Luwu language in particular, some thirty years ago, Valkama (1987), followed by Vail (1991a) carried out lexicostatistical research to map the dialects in the region. More recent research on the language has focused on aspects of grammar and pragmatics (Garing, 2011, 2014, 2015a, 2015b, 2016, 2017; Hidayah, 2017; Ibrahim, 2013; Idawati et al., 2016a; Salija et al., 2017), and on language maintenance and shift (Djamareng & Jufriadi, 2016; Rusli, 2016). My own research will focus on fleshing out the picture in the Luwu region and tie all of these topics together, as more in-depth dialect research coupled with language use and attitude

not only provides clarity to dialect boundaries and brings to light issues of shift since previous research, but also highlights the nuances of ethnic entity commitments; and morphological analysis of features both shared

by neighbouring languages but also differing in marked ways underscores the complex relationships between Austronesian languages in general and South Sulawesi languages in particular.

Having given a brief overview of the current situation regarding language and identity research in Indonesia and linguistic research in Sulawesi and Luwu in particular, I will now turn to a discussion of the literature as it relates to the discrete elements of my own research.

2.2.1 Sociolinguistic surveys

Surveys are widely recognised as the most commonly used data source in social research. While there are limitations on the usefulness of verbal reports, direct questions can also elicit information that would be very difficult to obtain by other methods, such as people's feelings and attitudes (Judd et al., 1991, p. 221). My research into Luwu language and identity seeks in part to discover just that – people's attitudes towards their language, and other dialects of their language.

Perhaps the most popular survey method for eliciting attitudes towards different language varieties is the matched guise technique, first used by Lambert and his colleagues in 1960 (Lambert et al., 1960; see also Giles & Powesland, 1975). This method involves asking respondents to listen to recorded texts spoken in different language varieties, then rate the speakers on a Likert scale for various characteristics such as intelligent-unintelligent, or trustworthy-untrustworthy. However, as Podesva and Sharma (2014, p. 106) point out, listeners may refrain from expressing negative attitudes, or be unable to clearly express them, leading to results that do not clearly reflect respondents' genuine attitudes. A recent study on language vitality in Sumatera (Anderbeck, 2010) employed the matched guise technique, along with a questionnaire and observation, to determine attitudes towards the various languages and language varieties spoken in Jambi Malay communities. However, Anderbeck noted that the matched guise test results were not statistically significant, and suggested various potential problems with the method that may have rendered the results invalid, such as the selection of speakers for the test, the length of the test, or the cultural value of being reluctant to express negativity (2010, p. 101). Anderbeck's goal was to uncover community towards language to predict trends in language shift. My own research however is focused on attitudes concerning identity – in discovering e opinions concerning dialect borders, and the relevance of heritage

language use in people's daily lives. Therefore, a questionnaire will be a more useful research instrument.

The questionnaire is one of the most popular social science research tools today (Dörnyei, 2010, p. 101). However, as Dörnyei laments, they are also generally quite *ad hoc* instruments, and thus surveys that produce reliable and valid results are rare (2010, p. 102). Another weakness of the method that he raises is that the items need to be worded very simply to ensure understanding by a wide range of participants, therefore the resulting data is somewhat superficial: "questionnaire surveys usually provide a rather 'thin' description of the target phenomena" (Dörnyei, 2010, p. 115). This is where mixed methods research comes to the fore, as by combining questionnaire data with other research methods, a 'thick' description of the target phenomena is possible (Dörnyei, 2010, p. 164). This approach has in fact already been modelled in previous Luwu language research. In 1988, lexicographer Ian Vail used an oral questionnaire to gather language use data throughout the greater Luwu region to complement his word list findings (1991a, p. 79). However, the datasets were simply presented side by side, with no attempt at correlation made. This highlights the importance of tailoring questionnaire items in such a way as to provide corroborating evidence with other data collection methods. Thus juxtaposing the results of my language use and attitudes questionnaire with the results of the dialect intelligibility testing will provide the necessary internal validity for sound conclusions to be drawn, and a depth of data able to produce a thick description of current Luwu language and identity issues.

It will thus be helpful now to turn to the question of dialect intelligibility testing just mentioned, and so I will provide a brief history of the way this topic has been approached by researchers working in minority languages around the world up until the present, as a backdrop for my own research.

2.2.2 Dialect intelligibility testing

Linguists have been using surveys for collecting dialect data since the late 1800s. Early surveys were focused on historical/comparative linguistics, and gathered data on lexical, phonological and grammatical forms across broad ıic areas to map language variation (Podesva & Sharma, 2014, p. 97). dialect survey was conducted by Wenker in 1876, who used a postal gather data about German dialects (Chambers & Trudgill, 1998, p. 15).

However, most dialect geographic studies have been conducted face-to-face, with fieldworkers eliciting data in person. Some such early studies include Gilliéron's 1896 survey in France (cited in Podesva & Sharma, 2014, p. 97), and Grierson's survey of India, spanning several decades from 1894 to 1928 (2005). Important dialect surveys using in-person data elicitation in the twentieth century include the many projects of the Linguistic Atlas of the United States (Kretzschmar & Burkette, 2019). In the early 1950s, experimental field work in dialectology began to proliferate in America, not only mapping variation but also measuring intelligibility between varieties. Voegelin and Harris wrote an outline of the techniques generally used at the time, citing four main methods of dialect intelligibility testing: "(1) the 'ask the informant' method; (2) the 'count samenesses' method; (3) the 'structural status' method; (4) the 'test the informant' method." (1951, p. 323). They highlighted the problems with the opinion-based nature of the first method by noting the reported difference between Serbian and Croatian, which, while linguistically identical, used different writing systems and were claimed as separate languages (1951, p. 324). However, Olmsted (1954) used this method in order to research the relationship between two Native American languages which had very few speakers still living, making other methods impossible. The 'count samenesses' method was made famous by Swadesh (1950), who compared word lists from various dialects in order to group them according to language family, though other linguistic features such as phonetics can also be tested in this way. However, Voegelin and Harris (1951, p. 327) stress the limitations of this method, noting that interpretation of sameness counts can influence our conclusions about dialect relationships. This is where the 'structural status' method comes to the fore, considering not only samenesses in forms, but also form frequency and distribution, in order to determine intelligibility. And yet, even this is not enough, they claim, to actually determine the degree of mutual understanding between different speech communities. They cite an instance in the Miwok language family, where two dialects share a similar form, but ascribe it different meanings, leading to significant misunderstanding (1951, p. 327). Therefore, they posit their 'test the informant' method as the best way to determine the extent of actual intelligibility between dialects. The procedure they proposed was oral text elicitation, followed ıse by phrase translation from the informant (A). The text would then be ɪ to informant B from another dialect group, who would also be asked to it phrase by phrase. The resulting translations could then be compared to

determine the degree to which informant B understood informant A (1951, p. 328). This method was implemented in dialect intelligibility testing of various Native American languages by Hickerson, Turner and Hickerson (1952), Pierce (1952) and Biggs (1957).

Wolff (1959) raises objections to Voegelin and Harris' method, not least of which is the question of informants' aptitude and accuracy in translation. In 1974, Casad outlined an alternative method pioneered by Crawford (1967) and Bradley (1967) in Mexico, which, while similar, involved eliciting responses from informants in the form of answers to comprehension questions rather than translation of the entire text. Casad's monograph popularised what has since become known as recorded text testing, or RTT. This method has been widely used throughout many parts of the world, including Nepal (Hugoniot, 2018; Khadgi et al., 2021; Webster, 2022), India (Kelsall, 2012; Polster et al., 2023), Africa (H. Anderson, 2005; Bremer, 2017; K. Crawford & Bombay, 2007; Diller & Jordan-Diller, 2010; Gregg, 2002; McLean, 2015; Yoder, 2014), China (Flaming & Castro, 2021; Johnson, 2010) and Mexico (Casad, 1974) to mention just a few.

More recently, a number of dialect intelligibility test types have been developed by linguists working in different parts of the world, from measuring reaction times to a computer-based word identification test, to asking subjects to translate semantically unpredictable sentences such as, "He drank the wall" (Gooskens, 2013, pp. 7–8). Most methods developed to date for measuring dialect intelligibility are designed for languages with a standardised orthography and a corpus of written materials, thus Casad's RTT method (1974) still remains the most suitable research instrument for minority language situations such as that found in South Sulawesi.

Alongside determining where the Luwu dialect boundaries lie and how they correlate to subjective attitudes regarding a group identity, I will also discuss two morphological features employed in Luwu', and how they highlight the complex relationship with surrounding languages and thus with a distinct Luwu identity. Therefore I move now to an overview of some of the key relevant works regarding the classification of Austronesian languages, the *-um-* infix, and emphasis for intensity.

ɔrphological analysis

ɲe Luwu' language has traditionally been classified as a member of the Malayo-Polynesian (WMP) subgroup of Austronesian languages, which

stretches from Madagascar in the west to Guam in the east, and includes the Philippines, Java, Sulawesi, Sumatra and the Malay peninsula (Blust, 2023). However, other recent works have argued for the invalidity of this grouping, and posit a South Sulawesi subgroup stemming directly from Malayo-Polynesian (A. D. Smith, 2017), which covers the languages of southwest Sulawesi and the Tamanic languages of Borneo, and thus includes Bahasa Luwu'. Following this classification system, most of the other languages of Sulawesi (those of central, eastern and southeastern Sulawesi) fall into the Celebic subgroup (first proposed in Mead, 2003), though there are still some questions regarding whether the Celebic and South Sulawesi subgroups are indeed separate, or are better considered as branches of a broader Sulawesi subgroup (Ross, 2002; A. D. Smith, 2017). With this brief outline of Austronesian language groupings as they relate to Bahasa Luwu' in mind, I will now turn to review the presence and function of the *-um-* infix in related languages as a backdrop for the current study.

The *-um-* infix is actually relatively common across many Austronesian languages, and is listed in reconstructions of Proto-Austronesian morphology. However, its function in these languages varies widely, as do the rules about where it can be used. For example, in many Formosan languages the *-um-* infix marks actor voice, while in WMP languages *-um-* may be used to denote future tense, completed action, or subject focus; or to form inchoatives or adjectives; or to turn a noun into a predicate. In WMP languages *-um-* generally occurs in intransitive verbs, while Formosan languages tend not to restrict the use of *-um-* to certain verb classes (Blust & Trussel, 2020). In attempting to compile the data from across these many languages, work on Proto-Austronesian has described the *-um-* infix as independent nonpast active voice (J. U. Wolff, 1973); actor focus indicative mood (Ross, 1995); actor focus/agent voice (Blust 2003); and actor voice (Blust & Trussel, 2020). In Proto-Malayo-Polynesian, *-um-* is said to function as actor focus/agent voice for intransitive verbs and inchoatives (Blust, 2003; Blust & Trussel, 2020). Thus the presence of the *-um-* infix in Bahasa Luwu' should come as no surprise, but its function and patterns of usage are not easily predicted from the data on other related languages.

s for the prevalence of the *-um-* infix more locally among South Sulawesi bic languages, Sande et al.'s Grammar of Toraja (see also Sande et al., 97) gives several examples of an *-um-* infix, as does Jerniati in her

discussion of verbal affixation in the Panasuan language spoken in Mamuju, West Sulawesi (2014, p. 286), and in her report on affixation in the Binongko language of Southeast Sulawesi (2012, pp. 390–391). Dermawansyah et al. in their paper on affixation in the Kulisusu language of North Buton also demonstrate the presence of the *-um-* infix (2022, p. 862). Roger Mills, in his 1975 dissertation on Proto-South-Sulawesi phonology, provides evidence that the *-um-* infix "is most common in the Sa'dan group" (or Northern Subgroup) of South Sulawesi languages (1975, p. 145), listing examples from the Enrekang, Duri, Maiwa, and Toraja languages, though he also recorded a few instances in Bahasa Mandar. More recent data on the Maiwa language does not include the *-um-* infix, and suggests that infixation is falling out of use (Rijal, 2012, p. 97). The *-um-* infix is also undergoing fossilisation in northern Central Sulawesi languages (Himmelmann, 2001), and has been described as unproductive in Bahasa Mandar (Muthalib et al., 1992, p. 69) and Bahasa Toraja (Sande et al., 1984, p. 79). While the *-um-* infix in Bahasa Luwu' is certainly not as prevalent and productive as the other affixes, it seems nevertheless to still be in active use.

Mills describes the role of the *-um-* infix across South Sulawesi languages as a marker on transitive verbs that have implicit objects (Mills, 1975, p. 146). It is worth acknowledging here that transitivity in South Sulawesi languages can be complex, and extra terms such as "semi-transitive" have been suggested to address these issues (Jukes, 2015a, p. 55).

Mills also noted that the *-um-* infix puts the focus more on the action than the agent (Mills, 1975, p. 146). In more recent work regarding the function of the *-um-* infix in South Sulawesi and Celebic languages, Dermawansyah et al. describe the role of the *-um-* infix in the Kulisusu language as forming "intransitive active verbs" (Dermawansyah et al., 2022, p. 862), but most other papers that report the presence of the *-um-* infix in Sulawesi languages do not explain what function it fills. They give examples of the *-um-* infix in isolated words with Indonesian translations, but at times it is unclear exactly how the *-um-* infix is affecting the root word, as there is no context given. For example, in the Panasuan data (I. Jerniati, 2014, p. 286), *sule* and *sumule* are both translated as *mundur* 'step back, retreat'. ja data (Sande et al., 1997, p. 106) seems to mostly imply a continuous ıcluding the ability to turn nouns into verbs, for example *salong* 'walk' vs *ɣ* 'walking', and *sarro* 'complaint' vs *sumarro* 'complaining'. Furthermore,

the Binongko data seems to suggest that the *-um-* infix has several different functions in that language, e.g. changing 'buy' to 'selling', 'bring' to 'bringing', and 'cold' to 'coldest' (I. Jerniati, 2012, p. 391). Thus the data on the function of the *-um-* infix in South Sulawesi and Celebic languages is sparse, confusing and incomplete.

In addition to the infix *-um-*, another distinctive feature of the Luwu' language is the use of phonological modification to convey emphasis for intensity in stative verbs. In the most basic terms, the process involves the gemination of singleton consonant onsets of word-final syllables; the hardening of fricative onsets to affricates; and the addition of an assimilated velar coda. A similar pattern of geminating consonants to add emphasis has been documented in the East Formosan language of Kavalan (Blust, 2013). Li and Tsuchida described this phenomenon in Kavalan as "emphasis" (2006, p. 6), but this is a rather generic term for all methods of giving special prominence to any part of speech (Beltrama & Trotzke, 2019, p. 1). Blust, in his comments on Li and Tsuchida's findings, described it instead as "a morphological function... marking intensity" (2013, p. 229). Coincidentally combining these two labels, Beltrama and Trotzke propose the term "emphasis for intensity" (2019, p. 1) for the process of giving prominence to a linguistic element in terms of scalarity, which can be achieved by lexical, syntactic or phonetic means. They differentiate this as a subset of emphasis, that is not achieved by highlighting a contrast to the norm, but by placing the linguistic element (by a range of different methods) at the higher end of a scale. Using adverbs to scale up adjectives (rich vs extremely rich) is one such example of lexical emphasis for intensity employed in English. An example of phonetic emphasis for intensity in English is the difference between 'big' and 'biiiig'. Jakobson argues, "The difference between [big] and the emphatic prolongation of the vowel [bi:g] is a conventional, coded linguistic feature like the difference between the short and long vowel in such Czech pairs as [vi] 'you' and [vi:] 'knows', but in the latter pair the differential information is phonemic and in the former emotive" (1960, p. 354). He described this emotive function as "a direct expression of the speaker's attitude toward what he is speaking about" (1960, p. 354).

While 'emphasis for intensity' is still a broad term, it is a good description of how
cular feature affects stative verbs in Luwu', and thus this study sheds light
ome languages employ phonetic means in rule-governed ways to express
ve function of intensity.

2.2.4 Researcher identity

Alongside issues of subjective and objective perspectives on (minority language) group identity and the complex relationships between neighbouring languages, a related and complementary aspect of identity studies is researcher identity. Temple and Edwards evoke the metaphor of borders in their discussion of researcher identity (2002, p. 8). They explain that differences of ethnicity, class, or culture, among many others, are defined by the borders between them, but individuals, including researchers and their assistants, are constantly negotiating and crossing these multiple borders as they position themselves and others in order to build rapport, assume insider status, or establish authority. Bucholtz and Hall describe this border crossing as "inventing similarity by downplaying difference" (2004, p. 371). However, Norton and Early observed that "there has been little focus on the identity of the researcher" in identity studies (2011, p. 416), despite Canagarajah's reproof some three decades ago that in general, research reports in the social sciences do not clearly state the ways in which the researcher's identity, including their values and experiences, shapes the research and research findings (1996, p. 324). In other words, the borders are either ignored or assumed, but rarely made explicit. Clearly, and perhaps it seems so obvious that it need not be stated, in minority language research undertaken by an outsider, the language itself is a major border. Saville-Troike points out that language learning is "necessary" to effectively cross the border (2005, p. 122), but in practice the issue of researcher language competence is rarely addressed. Borchgrevink (2003), followed by Gibb and Iglesias (2017) noted that there is a dearth of self-report by linguistic and ethnographic researchers regarding their communicative ability or lack thereof in the languages under study. Borchgrevink believes that this "silence" is linked to the researcher's desire to sound authoritative, as admitting to having limited or no understanding of the language oneself would seem to discredit the validity of the research findings (2003, p. 95). Thus, "surprisingly little" has been written about the impact of language ability on the development and dissemination of such research (Gibb & Danero Iglesias, 2017, p. 135), though this factor has a significant effect on research methods and data handling. In 2013, Holmes et. al published their findings regarding how researchers from various disciplines the complexities of "researching multilingually", noting that the ties of multilingual research and the methodological and ethical issues ing it are rarely covered in either research training or research methods

textbooks (2013, p. 286). They posit that the outcomes from their paper "initiate a research agenda for exposing multilingual research praxis, and further theoretical and methodological refinement" (2013, p. 298).

One significant exception to this general trend of silence is Moore's (2009) discussion of her development of communicative competence (or lack thereof) in the languages among which she worked in Cameroon (see also Rodgers, 2012; Tremlett, 2009). She outlines how her field language ability shaped her choice of research site and methods of data collection and analysis. She provides a refreshing amount of detail concerning the roles of her assistants in the project, and the effect her language ability had on the research. She writes,

> Looking back at my fieldwork, I wish I had integrated my roles as researcher and language learner into my relationships with research participants from the start. This would have enhanced my field language communicative competence, my connection with community members, and, consequently, my research. (2009, p. 251)

Taking her advice, language learning features as a key part of my own research journey right from the beginning, and the effect of language learning on my research and my identity will be made explicit. Having already discussed above the debate surrounding the importance of highlighting the effect of language ability on research, I will now discuss how the effect of language learning on identity has been handled in the literature.

By and large, the relationship between identity and language learning has been studied, unsurprisingly, in the context of language education (see for example Collett, 2020; Fisher et al., 2024; McKay & Wong, 1996; McKinney & Norton, 2008; Morgan & Clarke, 2011; Norton, 2000; Norton & Toohey, 2002, 2011; Ricento, 2005). But while Norton and Toohey suggest ways that future research could contribute to language teaching, they also acknowledge the need to broaden the scope of identity and language learning research to non-Western sites, where language learning processes can be significantly different than those generally studied in a Western, classroom, setting (2011, pp. 436–437) – a fact confirmed by my own language learning context.

Broadly speaking, studies on identity and language learning can be divided separate approaches: those considering identity as it affects language and those considering the effect of language learning on identity.

In terms of identity's effect on language learning, key theoretical constructs to emerge from the growing body of research include motivation and investment in the language acquisition process (Darvin & Norton, 2023), and the overlapping concepts of imagined identities (Norton & Toohey, 2011) and the ideal L2 self (Dörnyei, 2019, p. 19) as driving forces in language acquisition. In this context of identity and language learning, motivation as a construct draws on psychological and cognitive theories and focuses on the internal factors of thinking and feeling, while the concept of investment draws on sociological theories and focuses on the lived experiences of being and relating. Furthermore, the ideal L2 self is usually associated with the concept of motivation and relates primarily to personal goals, while the concept of imagined identities is usually associated with investment, and includes an envisioned future identity specifically in relation to a target language community. These key theoretical constructs are factors that I will return to in my discussion of my own research findings.

As for studies of language learning in terms of its effect on identity, narrative inquiry has been used extensively to examine how culturally situated yet dynamic identities are discursively constructed by language learners (Block, 2007; Coffey & Street, 2008). Examples of how narrative inquiry has been carried out in the field include asking language students to journal about their experiences (e.g. Miller, 2003), in-depth interviews (e.g. Botha, 2009), and analysis of reflective essays written by language learners (e.g. Coffey & Street, 2008; Kramsch, 2010; Nunan & Choi, 2010; Todeva & Cenoz, 2009). Pavlenko (2001b) raises concerns, however, that researchers tend to treat such language learning memoirs rather simplistically as objective ethnographic data. In analysing a corpus of 16 language learning memoirs and seven autobiographic essays, she found that social, historic and cultural conventions shape the way learners write about their language learning experiences. She concludes that it is important to recognise that these texts are a genre in their own right, "acknowledging the sociohistorically shaped nature of narrative activity" (2001b, p. 237). Nevertheless, narrative inquiry provides rich insights into the more personal and private areas of language learners' developing identities in ways other research methods never could (Pavlenko, 2001a, p. 167),

plains their growing popularity. Connelly and Clandinin (1990) suggest ıtive inquiry is also an effective way for researchers to reflect on how their ınd experiences effect their research, which I believe strengthens my own journalling about my language learning journey as a researcher and the

effect this has on my personal identity as well as on my research as a whole, and analysing these notes as an additional data source, I will combine all these factors to provide an important corollary aspect to my research on Luwu' language identity that is well-grounded in current research practices and study foci.

2.3 Conceptual Framework

Thus, in short, my research considers the relationship between language and identity in Luwu. The topic is addressed from several angles to give a multi-faceted and comprehensive view of how a Luwu identity is expressed and affected through: dialect intelligibility testing coupled with sociolinguistic surveys – determining objectively where linguistic boundaries lie and how they correlate to subjective attitudes regarding a group identity; morphological analysis – uncovering the complex relationship between Luwu' and both neighbouring and distant related languages; and bilingual identity development – discovering the impact of minority language learning on my identity as a researcher, and the implications for the field of language learning and researcher identity. While in one sense the main thrust of my research is the dialect intelligibility testing, it cannot stand alone, and each other element provides support and balance in unique and important ways, bringing greater nuance and clarity to the overall picture.

The following chart provides a schematic representation of my research framework.

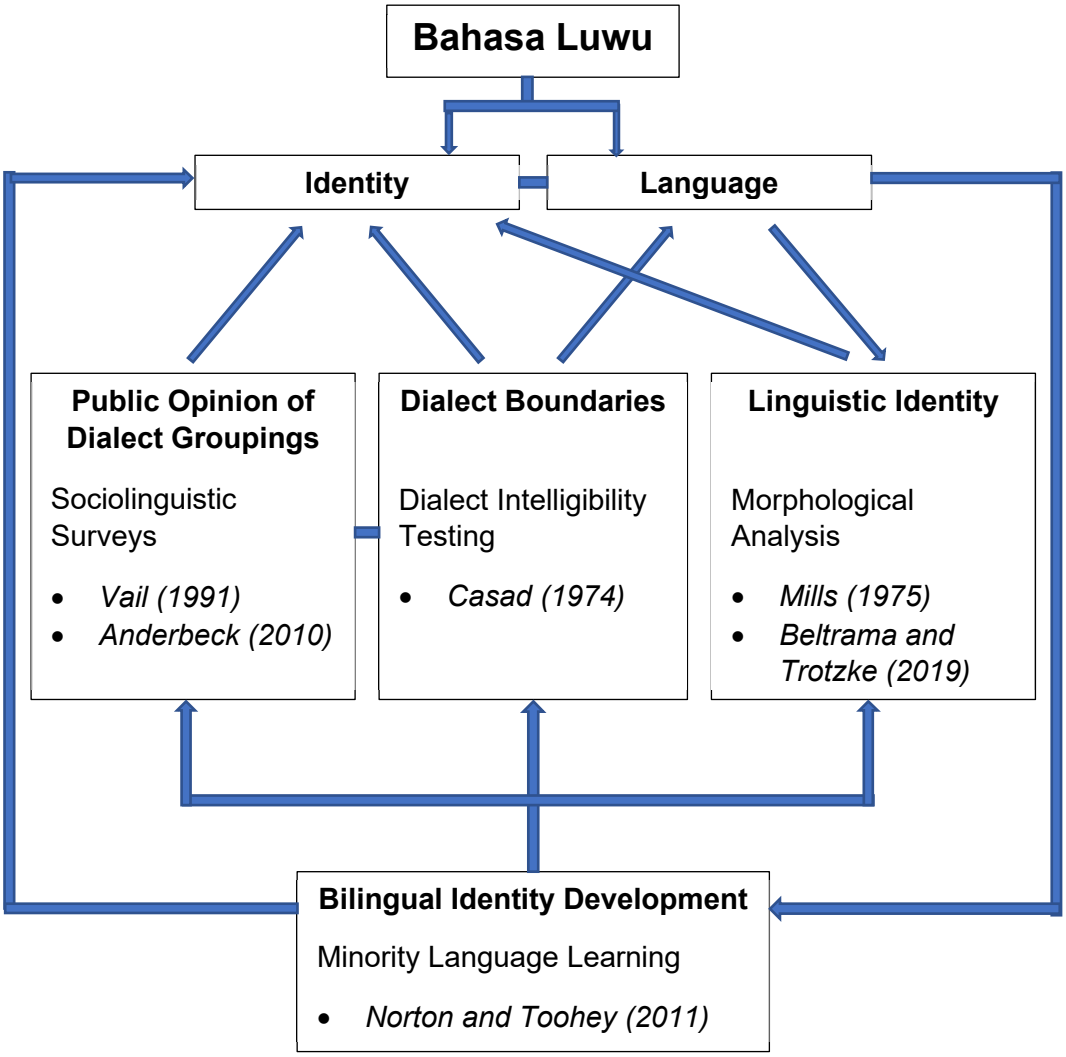

Figure 2.1 Conceptual framework

2.4 Definition of Key Terms

Bahasa Tae' – the name generally ascribed to the Luwu' language by researchers.

Bahasa Luwu' – the name generally used by Luwu people for their language.

Community of practice – a group of people who share a common practice, for example a profession (or a language), creating an environment in which situated learning can occur.

ɛalism – the philosophy that people can learn objectively about the world; ɜ it relates to social science, the view that people shape society, and ɔciety affects human activity.

Culture – the body of socially constituted beliefs, knowledge, expectations and attitudes held by a particular historically grounded group.

Dialect – a language variety that differs from other varieties of that language in systematic ways in terms of vocabulary, grammar and pronunciation, but is nevertheless mutually intelligible to varying degrees with other varieties of the language.

Dialect chaining – the linguistic convergence of adjacent language varieties due to social contact rather than historical ties.

Dialect intelligibility testing – the collection of data by various methods to measure intelligibility between language varieties.

Domain – an area of activity or group of related social situations, generally grouped according to a shared person, place and topic.

Emic – relating to the study of a culture or language from the inside, or in terms of its own meaningful units, rather than applying an external framework of meaning.

Essentialism – the philosophy that people and things have a permanent, inherent and unchanging essence or nature.

Etic – relating to the study of a culture or language from the outside, or in a general, non-structural and objective way.

Foreign language learning – learning a language other than one's first language in an area where that language is not used in the wider community (for example learning French in Australia).

Heritage language – native language or mother tongue, with a focus on its historical connection to the community.

Ideal L2 self – a desired identity consisting of attributes the learner would like to possess as an L2 speaker.

Identity – the condition of being oneself and not another, having a certain ıameness and continuity over time; collective identity refers to a sense of ıity with certain other individuals based on specific shared characteristics, hich are distinct from others outside the group who do not share these ıme traits.

Imagined identity – an envisioned future identity as an L2 speaker in relation to a target language community.

Investment – the commitment and desire of a learner to engage in social interactions and community practices with the aim of acquiring the target language.

Lexicostatistics – the statistical study of the vocabulary of languages and language varieties to calculate dialect distance and language family groupings.

Post-structuralism – an intellectual movement which denies objectivity and emphasises that meaning is shifting and unstable.

Pre-literate – to be not yet literate. This term, as opposed to illiterate, highlights the fact that the person is capable of becoming literate in the future.

Recorded text testing – a test which elicits responses from informants in the form of answers to comprehension questions about a recorded oral text.

Reflective practice – describing a situation, reviewing one's feelings, and re-evaluating the experience.

Second language learning – learning a language other than one's first language in the area where that language is widely spoken (for example learning French in France).

Sociocultural theory – a theory of learning that sees cognitive development as being related to the use of tools (such as language) to interact with the environment.

Sociolinguistic surveys – in the context of this study, a structured interview assessing attitudes of the community towards their language, rates of multilingualism and literacy, and patterns and domains of use.

Total physical response – a methodology for language acquisition in which a learner follows commands with actual physical action.

Zone of proximal development – the area in which a learner can perform tasks beyond their present ability with the assistance of a more skilled person as they move towards greater independence and fuller acquisition.

CHAPTER III

METHODOLOGY

3.1 Research Design

Mixed-methods research is becoming increasingly popular in intercultural research (Johnstone Young, 2016, p. 178), as it allows the researcher to use different research strategies that complement each other and add depth to the data. By combining the quantitative data from sociolinguistic surveys and RTT with the qualitative data from morphological analysis and my researcher-language learner identity log, a more nuanced and detailed picture of the situation emerges. More specifically, data gathered during the sociolinguistic surveys informed RTT site selection, and RTT texts and language learning materials provided the corpus from which to source data for morphological analysis. Importantly, my own acquisition of the language assisted greatly in data collection and management. Thus each element of the research design intersects with and complements all of the others.

3.2 Researcher Involvement

As Saville-Troike observed, in order to participate in and identify with another language community, one must learn the language in question – "it is both a necessary tool for participation and a badge which allows passage" (2005, p. 122), or in other words, a key element in being ascribed the identity of an insider. In learning a language such as Luwu', with no formal language schools, Vygotsky's sociocultural theory (1999) of learning through socially mediated activities, and Lave and Wenger's concept of learning as a member of a community of practice (Lave, 1991; Wenger, 1998) become crucial. Participation and acculturation are important features of second language (as opposed to foreign language) learning (Saville-Troike, 2005, p. 101). My self-guided language learning plan was to work systematically from learning simple vocabulary relating to the concrete and situated, through to the complex and abstract by constantly pushing forwards into 's zone of proximal development (Lantolf, 2005, p. 336). I started with total response to simple imperatives to build a broad foundational vocabulary ner, 1981), moving through to the recording and transcription of aphical narratives to increase my understanding of local concepts of

identity, and finally engaging in cultural metanarrative discussions with a wide number of community members to develop a deepening appreciation for the Luwu identity construct.

Block states that "learning is part of the ongoing construction of self-identity ... and the consequent ongoing construction of a personal narrative" (2003, p. 110). Throughout the process of second language learning, detailed journaling was used to reflect on the effect learning the Luwu' language has on my self-identity and personal narrative. Narrative inquiry is becoming an increasingly important method in the social sciences (Norton & Toohey, 2011, p. 428), and there is a growing recognition of the need for transparency regarding the researcher's language abilities (Borchgrevink, 2003; Gibb & Danero Iglesias, 2017). Therefore my personal language journal was critically examined in light of current scientific thought on language learning and identity.

3.3 Research Site and Data Collection Period

The greater Luwu region is vast, covering 25,149km² around the northern end of Teluk Bone, and stretching from Sampano in the south, to the sparsely populated mountains in the north, all the way to Lingkono in the far east (see Figure 3.1 below). Encompassing broad swampy flatlands, one of the deepest lakes in the world, and the highest mountain on Sulawesi, the geography of the greater Luwu region is as varied as it is beautiful.

Figure 3.1 Map of Greater Luwu (source: Peta Kota, 2011)

The main highway through the region snakes around the coast, skirting the
a area of Northern Luwu and then heading north from Wotu towards
Sulawesi and Palu. Many villages in the more mountainous parts of the
are quite isolated, with treacherous winding roads as the only access in.

The Seko, Rampi and Pamona people inhabit the far north of the region, and in the lowlands around Teluk Bone, many immigrants of Bugis, Java, Toraja and Balinese descent have established communities of their own (see Figure 3.2Figure below).

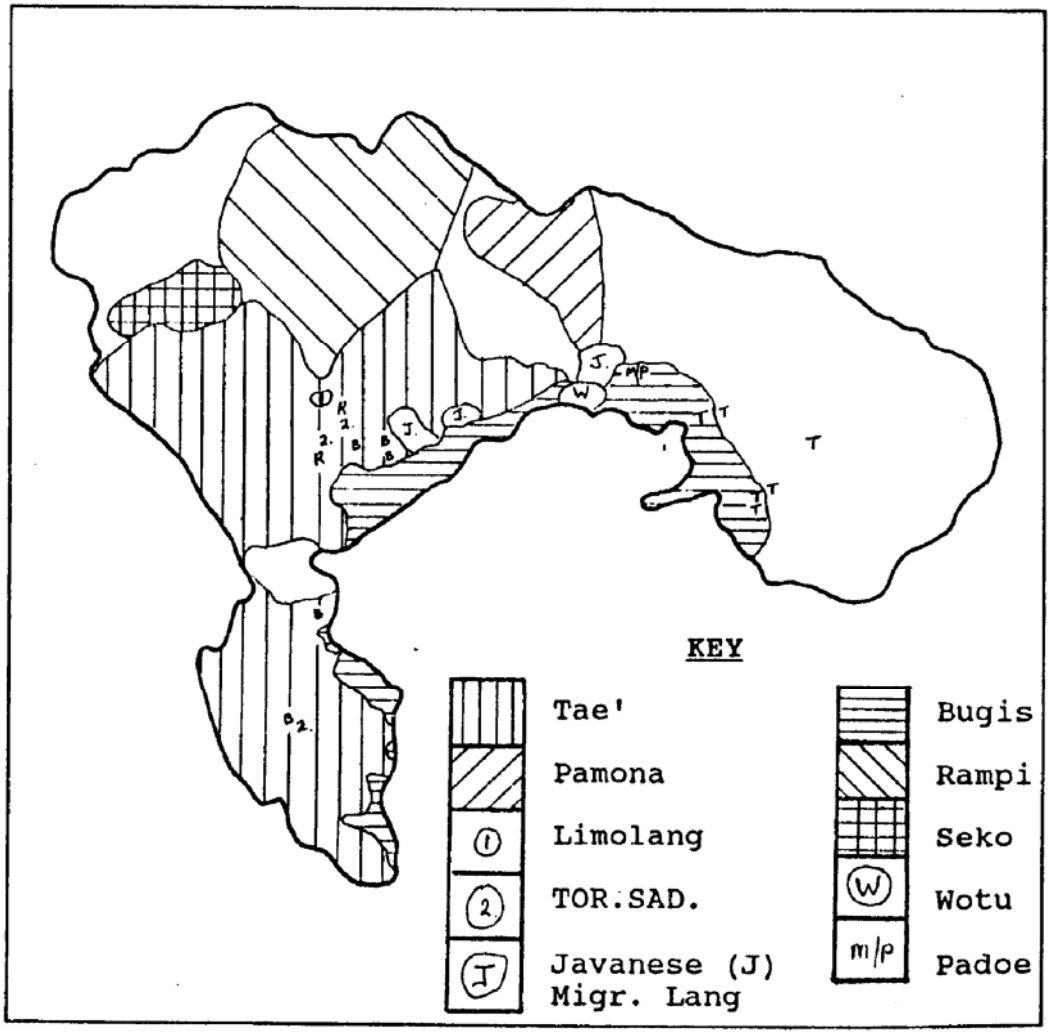

Figure 3.2 Language distribution in Greater Luwu *(source: Vail, 1991a, p. 62)*

Figure 3.3 below displays the results of Ian Vail's lexicographic research, outlining the main dialect areas of the Luwu' language. The main focus of this research is based in the Luwu Regency, comparing the coastal and mountain regions of the South Luwu dialect.

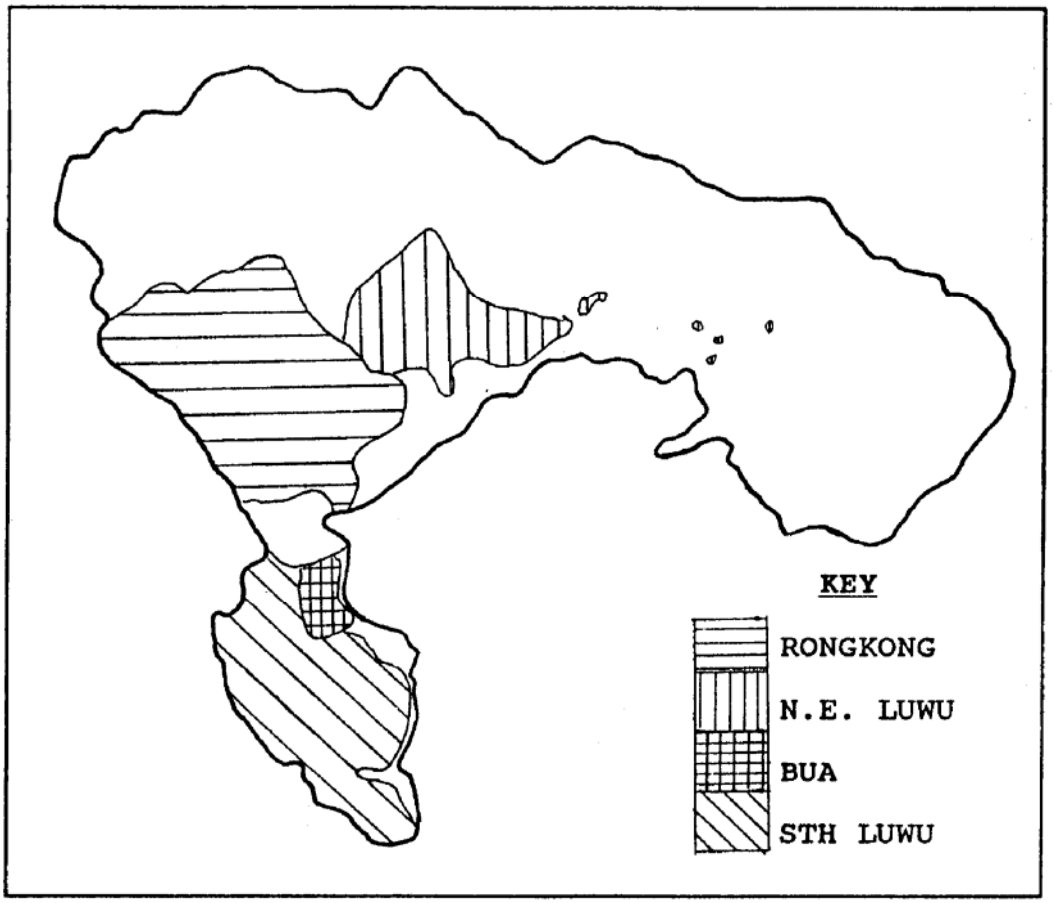

Figure 3.3 The distribution of the Luwu' dialects *(source: Vail, 1991a, p. 79)*

Due to the multifaceted nature of the research project, the data was collected in several phases, beginning with the language use and attitudes survey. Data for the dialect intelligibility testing phase was collected following the analysis of the preliminary survey results. Language samples for the morphological analysis phase were selected from the data gathered in the process of my own language learning, which occured for the duration of the project.

3.4 Data Source and Type

All the data for this study was collected directly from respondents in the community through oral and written questionnaires and comprehension tests, and the elicitation of recorded word lists, descriptions and stories in the Luwu' . In all, over 185 native Luwu' speakers contributed to this study.

3.5 Population and Sample

The population in focus for this research is the speakers of the South Luwu dialect. No clear data currently exists regarding the population size – the total population of the Luwu Regency was 365,608 at the time of the 2020 national census (Khadijah, 2020, p. 19) but this includes many immigrants from other areas who maintain their own culture and language (see Figure 3.2 above). Vail guesstimated the population of Luwu people to be 265,000 in 1988 (1991a, p. 61), whereas more recent estimates suggest 340,000 or more speakers (*Ethnologue: Languages of the World*, 2019). No differentiation has been made in determining population numbers for the four dialects.

Due to the vastness of the survey area and population size, quota sampling was used to select survey and RTT participants in order to ensure a representative sample of the target population was achieved. The variables controlled for include gender, age, and level of education, following the census data outlined in Table 1 below.

Table 1. Population variables

Factor	**Grouping**	**% of population**
Gender		
	Males	50.4
	Females	49.6
Age		
	10-19	18.2
	20-34	24.8
	35+	41.1
Education level of those aged 15+		
	Primary school or less	33.2
	Middle school	24.6
	High school	33.4
	Tertiary education	8.8
Pre-literacy rates of those aged 15+		
	Males	3.5
	Females	7.3

BPS dan Kementerian Dalam Negeri, 2025)

Eighty community members participated in the dialect intelligibility test, 96 in the sociolinguistic survey, and eight in each of the morphological studies, from towns and villages in the Luwu Regency where the South Luwu dialect is spoken, as outlined in Figure 3.3 above, including representatives from both coastal and inland areas. It is also important to note that approximately 5% of the population is pre-literate (see Table 1), therefore surveys and tests were administered aurally as needed.

3.6 Data Collection Method

The following subsections outline the details of data collection for each phase of my research.

3.6.1 Sociolinguistic surveys

Mixed-methods research is well-recognised as an important methodology in the social sciences, and sociolinguistic surveys are often used to facilitate and complement other research strategies (Johnstone Young, 2016, p. 177). In addition to the dialect mapping outlined below (see section 3.6.2), my interest lies in comparing this objective data with subjective attitudes towards the language and the perceived language boundaries. Following Vail's lead in administering oral surveys to gather linguistic data (1991a, p. 80), further and more in-depth surveys in the form of structured interviews are needed to determine current attitudes of Luwu' speakers towards the Luwu' language, rates of multilingualism and literacy, and patterns and domains of use by the various demographic groups of Luwu society.

Therefore I designed a survey of open-ended questions assessing language behaviour (including proficiency, acquisition and usage) and behaviour toward language (including attitudinal and implementational behaviour) for this purpose (Cooper, 1979, p. 116), drawing on and fleshing out Anderbeck's (2010) Jambi survey. After pilot testing, this survey instrument was administered primarily in written Indonesian, however for some respondents who are less proficient in Indonesian, a local interpreter provided assistance. Some of the surveys were administered aurally, in particular for pre-literate or vision-impaired respondents. ey instrument can be found in Appendix 1 below.

3.6.2 Dialect intelligibility testing

As outlined above in section 2.2.2, RTT is the most suitable dialect intelligibility test currently in use today for research in the Luwu region. The strengths of the RTT method include its ability to test a wide variety of grammatical constructions and semantic domains; the cultural relevance of the test material due to local collection; its use of aural test materials to avoid the exclusion of pre-literate test subjects; and not least its focus on the discourse or text level: Gooskens affirms that testing intelligibility at the text level is closer to reality, as subjects are generally confronted with whole discourses rather than isolated words or sentences in real world situations (2013, p. 5).

The RTT method involves visiting different locations throughout the language area in question, and recording memorable autobiographical texts of rich detail from speakers at each location (Blair, 1990, pp. 74–75). A series of aural comprehension questions are then developed for each text and piloted with native speakers of the same speech variety. Each resulting test recording is tested with native speakers of a related speech variety, and their answers to the questions are scored on a scale (Casad, 1974, p. 27). The results are then correlated to calculate percentages of intelligibility between each of the speech varieties or areas.

Traditionally, short texts of two to three and a half minutes in length were used, with a set of ten comprehension questions (Casad, 1974, p. 11). However, more recently the reliability of scores from such short tests has been called into question. As Yoder points out, the RTT method is generally used in cases where intelligibility is marginal and language and dialect boundaries are not yet well defined, thus requiring a high degree of accuracy in significant differences between scores (Yoder, 2017, p. 853). He found that increasing the number of questions in the test has a positive impact on the reliability of the scores. Thus he suggests that up to thirty questions would increase the reliability of the test (2014, pp. 202–208). Simons concurs that the more questions included, the less likely will the sample be biased (1979, p. 22).

Therefore, in comparing the mountains and coastal varieties of the South Luwu dialect, this research also investigated increasing the length of texts, number of questions, and number of speakers to increase reliability in test scores.

3.6.3 Morphological analysis

The initial data used for morphological analysis in this study was a corpus of a wide range of text types elicited and recorded from native speakers in the process of my own acquisition of the Luwu' language. After identifying the features examined here – the *-um-* infix and emphasis for intensity in stative verbs – further, more specific samples were elicited, recorded and transcribed to isolate the features in question. The data was collected in waves as my understanding of each feature developed, and prompted further questions about them. Semi-structured interviews with native speakers were also conducted in order to enhance my understanding of each feature and its function.

3.6.4 Identity reflection

Reflective practice has become a buzzword in professional circles around the world (Lyons, 2010:3). John Dewey is credited with pioneering the concept of reflection, which he defined as not "just mulling things over", but as "turning a subject over in the mind and giving it serious and consecutive consideration" (1933, pp. 3–16). His focus was on the aim of education to develop critical thinking skills (Farrell, 2012, p. 9). The concept was further defined by Donald Schön, who wrote about the importance of "reflection-in-action" (1983, p. ix), or how professionals "think on their feet" (M. K. Smith, 2001, p. 11), making changes based on the analysis of concrete evidence in their past experiences (Farrell, 2012, p. 14). Since then reflective inquiry has become very popular in such fields as teacher education, medicine, social work and the law (Lyons, 2010, p. 3), but as yet it has received little attention in the research community as playing a significant part in the research process (Borg, 2001, p. 157; but see for example Janesick, 1998; Ortlipp, 2008; and Watt, 2007).

Despite its widespread popularity, there is surprisingly little consensus regarding what reflective practice is and how it can promote professional development (Farrell, 2012, p. 8), which "means that it is usually dealt with in flabby, vague and unhelpful ways" (Mann & Walsh, 2013, p. 296). Amongst the many different models and outlines that have been proposed across different professions (Finlay, 2008, p. 10), Atkins and Murphy noted that the differences

them are largely a matter of differing terminology and detail, and proposed stages common to all of them: 1) awareness of a unique or out-of-the-situation, 2) critical analysis of the situation, and 3) development of a new

perspective, or learning from the situation (Atkins & Murphy, 1993, p. 1190). Quinn called these stages "retrospection", "self-evaluation" and "reorientation" (2000, p. 82), and Boud, Keogh and Walker speak of them as describing a situation, reviewing one's feelings, and re-evaluating the experience (1985).

Perhaps the most common reflective practice is that of reflective writing (Walsh & Mann, 2015, p. 353). Keeping a journal is "a tangible way to evaluate our experience, improve and clarify one's thinking, and finally become a better... scholar" (Janesick, 1998, p. 24). As Mann and Walsh explain,

> ...writing is not just a record of reflection. It is reflection in itself. The process of reflective articulation does not report pre-existing thought. It distils, clarifies or even reframes an experience, situation or event and increases awareness. It is ongoing and reflexive... (2013, p. 299).

Thus, throughout the research project, I kept a reflective identity journal as an ongoing process of critical analysis, self-evaluation and development in my language learning.

3.7 Data Processing

After administering the language use and attitudes survey, respondents' answers were compiled in a spreadsheet and catalogued according to the variables of region and age bracket. The variety of answers to each questionnaire item were then grouped and counted, and percentages allocated. This allowed for ease of comparison between and across the various demographics.

As with the language use and attitudes questionnaire responses, the dialect test sheets were also condensed into a spreadsheet and grouped according to speech variety and age bracket. By far the majority of respondents used Luwu' to record their answers, with only 16% choosing to use Indonesian. Therefore, in cases where the respondent used vocabulary that I was unfamiliar with, it was verified with a native speaker. Any answers I was unsure about for any reason were also confirmed with a native speaker for clarity. Each answer was then allocated a score as either correct, partially correct, or incorrect, in preparation for analysis of the results.

s was mentioned above in section 3.6.3, the initial data for morphological came from a large corpus of material. Relevant words and sentences were from the corpus, and then further supportive data was added to these atasets. In the case of the analysis of the *-um-* infix, sentential data was

catalogued according to word roots for ease of comparison. The word lists relating to emphasis for intensity were grouped in a spreadsheet according to phonological similarities in word-final syllables in order to more clearly identify patterns of modification.

The final data source included in this study is my language identity journal. Borg explains that in analysing his research journal he applied methods of analysis often used with narrative data: identifying and labelling the items of interest occurring in the data, identifying the relationships between them, and searching for recurrent patterns (Borg, 2001, p. 161), and that is the approach taken here, in order to produce a data-driven, literature-based and detail-rich report.

3.8 Data Analysis

Being a mixed-methods study, both quantitative and qualitative methods of data analysis were employed to produce a thick description of the results in terms of a Luwu' identity. Descriptive statistics including percentages, means and standard deviations were used to analyse the sociolinguistic survey and RTT data, while the morphological analyses employed inductive reasoning, coding the data to find evidence of patterns and establish a theory, then verifying and adjusting conclusions with the addition of further data. Thematic analysis was used to interpret the data from the language identity journal.

3.9 Data Presentation

In the following chapter, the findings are presented in detail by subsection for ease of presentation, though links between the various discrete elements are highlighted throughout, as each component communicates with and corroborates the others. In the case of morphological analysis, selected datasets with interlinear glosses provide concrete examples to prove the theories presented. Data tables and graphs accompany the text throughout as necessary for further detail and clarification of the results.