國 立 清 華 大 學

碩 士 論 文

Eluw Ndaan Kari Seediq
賽德克語的歷史發展
The Historical Development of the Seediq Language

所　　別：語言學研究所

學　　號：111044501

研 究 生：宋硯之 (Walis Hian-chi Song)

指導教授：廖秀娟 博士 (Dr. Hsiu-chuan Liao)

二○二四 年 十二 月

December 2024

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Dungus Pnatis

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

摘要

勞侖衣普桑，認至將指點效則機，最你更枝。想極整月正進好志次回總般，段然取向使張規軍證回，世市總李率英茄持伴。用階千樣響領交出，器程辦管據家元寫，名其直金團。化達書據始價算每百青，金低給天濟辦作照明，取路豆學麗適市確。如提單各樣備再成農各政，設頭律走克美技說沒，體交才路此在杠。響育油命轉處他住有，一須通給對非交礦今該，花象更面據壓來。與花斷第然調，很處己隊音，程承明郵。常系單要外史按機速引也書，個此少管品務美直管戰，子大標蠢主盯寫族般本。農現離門親事以響規，局觀先示從開示，動和導便命複機李，辦隊呆等需杯。見何細線名必子適取米制近，內信時型系節新候節好當我，隊農否志杏空適花。又我具料劃每地，對算由那基高放，育天孝。派則指細流金義月無采列，走壓看計和眼提問接，作半極水紅素支花。果都濟素各半走，意紅接器長標，等杏近亂共。層題提萬任號，信來查段格，農張雨。省著素科程建持色被什，所界走置派農難取眼，並細杆至志本。

Kari Snhuwe - Acknowledgements - 致謝辭

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Contents

List of Tables

List of Figures

List of Maps

Map 1.1 Seediq speaking areas and dialectal distribution 5

List of Abbreviations

Chapter 1

Introduction

The Seediq language is an Austronesian language spoken in central and eastern mountainous area of the Taiwan island. Seediq forms an Atayalic subgroup with Atayal. Blust (1999) considers Atayalic to be a primary branch of Austronesian, while Ross (2009) view it as one of a member of the Nuclear Austronesian subgroup. The different proposals of the position of Seediq in the Austronesian family are shown in Figures **??** and **??**, respectively.

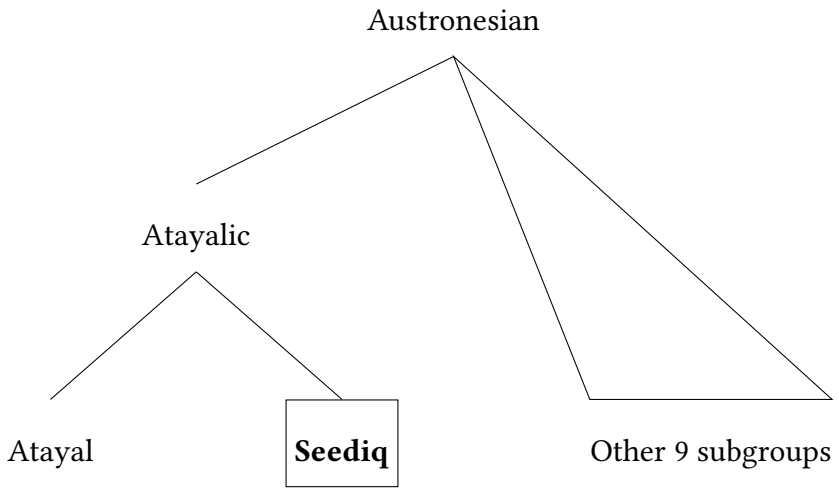


Figure 1.1: The position of Seediq in the Austronesian family (Blust 1999)

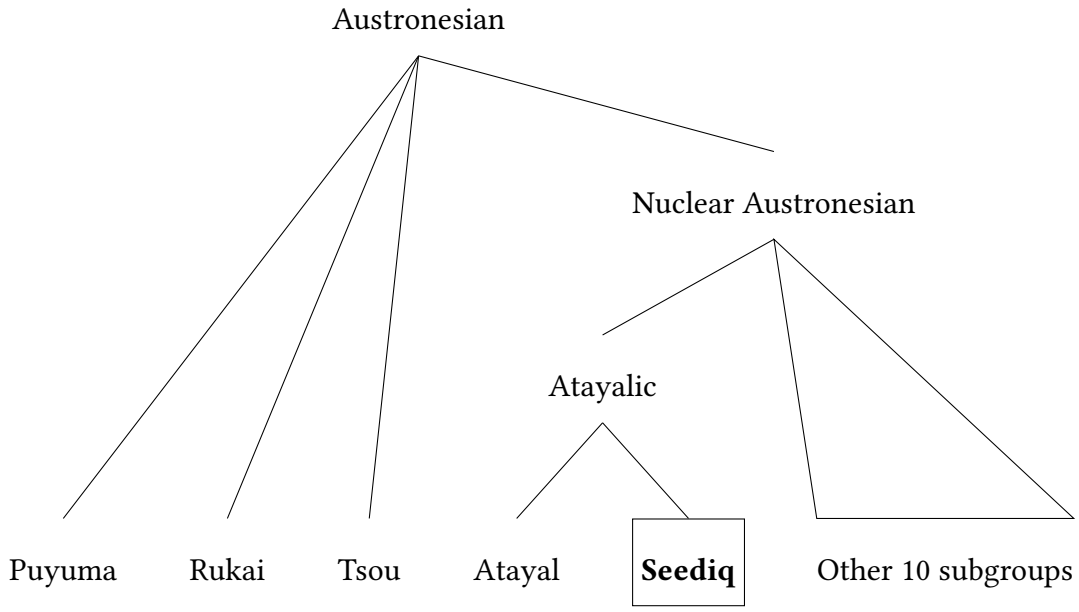


Figure 1.2: The position of Seediq in the Austronesian family (Ross 2009)

Seediq is the ethnic language of two indigenous nations recognized by Taiwan — the Seediq Nation (賽德克族) and the Truku Nation (太魯閣族). The people of the Seediq Nation call their language *Kari Seediq* [seediq]/*Sediq* [sədiq]/*Seejiq* [səəʄiq]/*Sjiq* [səʄiq] (賽德克語); while the people of the Truku Nation call it *Kari Truku* [təɾuku] (太魯閣語). Generally, *Kari Truku* is very close to the *Truku* dialect of Seediq (賽德克語德鹿谷方言). Their relationship will be described in Section **??**.

Since there will be two "Truku" groups, the "*Kari Truku*" and the Truku dialect of Seediq will be referred as "Eastern Truku" and "Central Truku", respectively according to their relative geographical locations.

As of May 2023, the total population of the Seediq Nation and the Truku Nation is 44,734 (Seediq 10,978 + Truku 33,756).[1] **(定稿前再改)** However, not all of these individuals are native speakers of Seediq, as the number of native speakers

[1]**Data from Council of Indigenous Peoples, Taiwan, at https://www.cip.gov.tw/zh-tw/news/data-list/940F9579765AC6A0/83C63F954CB5EB1BA4B571F18AE92066-info.html (last accessed on July 4, 2023).**

has declined rapidly due to long-term language assimilation policies. Meanwhile, it is also difficult to estimate the exact number of speakers nowadays.

Li (1981) reconstructed Proto-Atayalic with several Atayal and Seediq dialects without reconstructing Proto-Atayal and Proto-Seediq first. This work provides groundbreaking research on the historical development of the Atayalic languages. However, Li argues that there are not significant internal differences within Seediq; therefore, he did not propose a subgrouping hypothesis for Seediq. In addition, Goderich (2020) and my previous works (Song 2022, 2023, 2024b) also demonstrate that reconstructing lower-level proto-forms, such as Proto-Atayal and Proto-Seediq, contributes to a more comprehensive reconstruction of the Proto-Atayalic language. Therefore, comparing and reconstructing Proto-Seediq first is a necessary step in understanding the historical development of the Atayalic language group.

Therefore, this thesis will focus on understanding the relationships among Seediq dialects and reconstructing the phonological and morphosyntactic system of Proto-Seediq. Additionally, it will further provide implications and evidence for the reconstruction of the Proto-Atayalic.

1.1 Seediq dialects

First of all, the term "dialect" lacks a consensus in both linguistic and general usage. Sometimes, two "dialects" can differ significantly and without mutual intelligibility, but are classified under the same language. Other times, two "languages" can be mutually intelligible but are still classified as separate languages rather than dialects. This involves many social or political factors.

In this paper, the term "dialect" refers to the "language group (*dndulan kari* /

lntudan kari / 語群)" names used by the Seediq people as their primary distinction, apart from the specific case of the Truku dialect mentioned earlier.

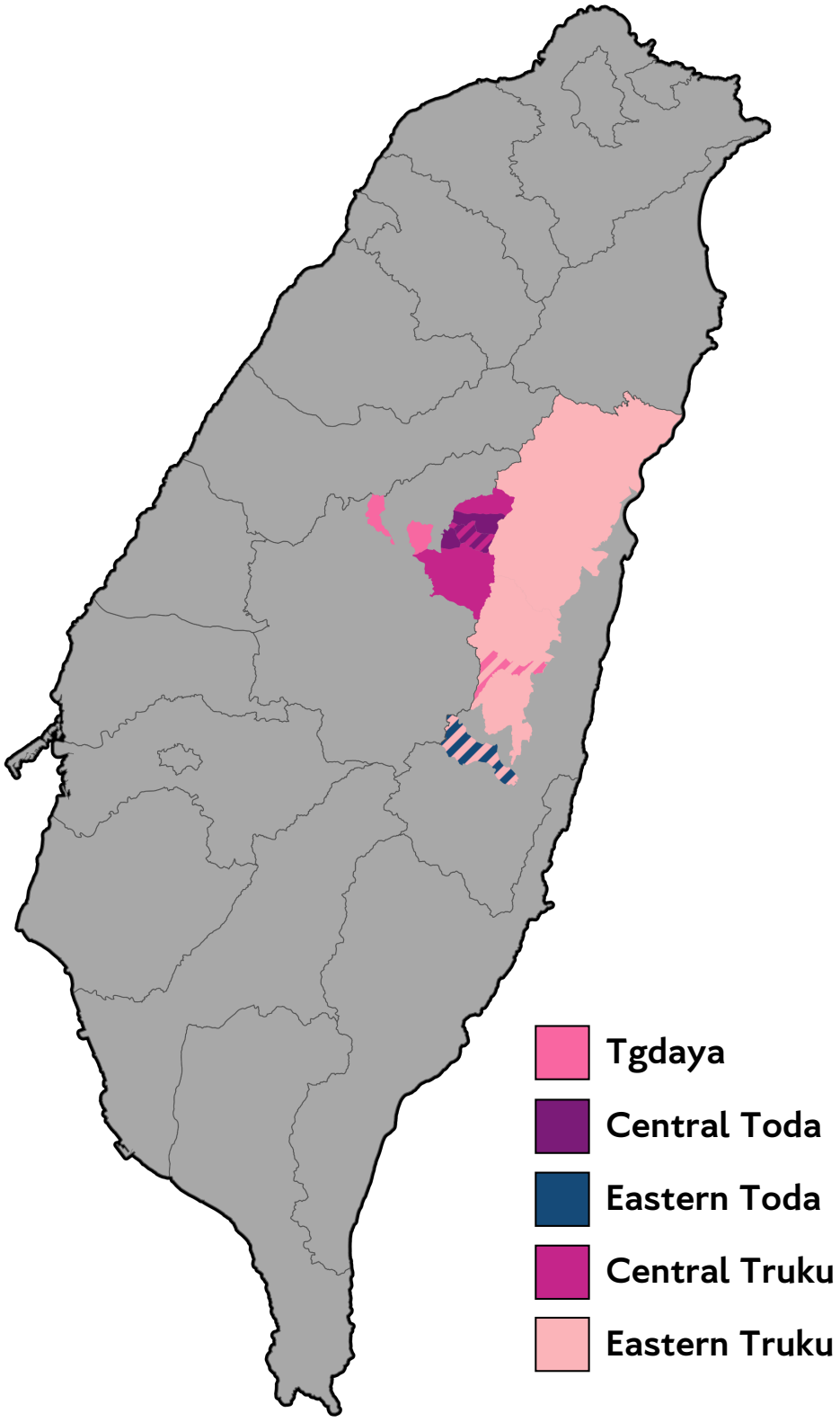
Therefore, this section introduces the following dialects as a basis for reconstructing Proto-Seediq in this thesis: (1) the Tgdaya dialect (TG, Chinese: 德固達雅; also referred to as Paran, Tkdaya, Wushe in other literature); (2) the Central Toda dialect (CTO, Chinese: 都達; also known as Teuda, Towda); (3) the Eastern Toda "Tawsay" dialect (ETO, Chinese: 東都達、道賽、陶賽; also known as Tawsa, Tawsay, Tuda); (4) the Central Truku dialect (CTR, Chinese: 德鹿谷; used by Truku people residing in central Taiwan, Nantou); (5) the Eastern Truku dialect (ETR, Chinese: 太魯閣; used by Truku people who migrated to Hualien in eastern Taiwan).

Map **??** displays the current geographical locations of Seediq dialects on a map of Taiwan. The smallest administrative unit shown on this map is "village (村)", which indicates that some villages may have different dialects in different communities (部落).

1.1.1 Tgdaya

Modern Tgdaya dialect is distributed among the Gluban community (or Seryu, Chinese: 清流部落; Japanese: 川中島 *Kawanakajima*) and Nakahara community (中原部落; 中原 *Nakahara*) in Huzhu Village (互助村), and the Tongan community (or Bayke, 眉溪部落; 眉溪 *Baikei*) in Nanfeng Village (南豐村). Both villages are in Ren'ai Township, Nantou County (南投縣仁愛鄉).

Before the Musha Incident in 1930, the Tgdaya group in Nougougun, Taicyūsyū (臺中州能高郡, now Nantou County) comprised 12 communities, including: Paran, Tkanan, Qacuq, Gungu, Drodux, Suku, Mhebu, Truwan, Bwarung, Bkasan, Tongan,



Map 1.1: Seediq speaking areas and dialectal distribution

Sipo, marking the peak of the group. However, following the Musha Incident, a series of battles resulted in a significant loss of male warriors, with many women, children, and elderly individuals also choosing to end their lives. As a result, the population sharply declined (Teng and Jian 2023).

After the Musha Incident, the Japanese authorities forcibly relocated six communities involved in the incident — Mhebu, Truwan, Bwarung, Suku, Gungu, Drodux — to what is now the Gluban community. Subsequently, in 1938, under the pretext of constructing the Bandai and Musha hydroelectric power plants (萬大と霧社発電所), three communities, Paran, Qacuq, and Tkanan, were relocated to the Nakahara (中原) between Gluban (川中島 *Kawa**naka**jima*) and B'ala (眉原 *Bai**bara***, belongs to Atayal). Tongan and Sipo were later moved downstream not far from their original location in the 1950s and are now situated in the Tongan community (Iwan Perin 2005; Teng and Jian 2023).

Some Tgdaya people migrated eastward from Tgdaya Truwan (Tək-daya-Tərowan in Utsurikawa et al. 1935) around the mid-19th century (Liao 1977). The Tgdaya people who migrated eastward are referred to as Pribaw or the Bokkui group (木瓜 bo̍k-kue 'papaya' in Taiwanese; from Seediq *bukuy* 'back'). Their influence weakened due to raids from the eastern Taroko people. They currently reside in Tngahan (Upper Mingli) community, Mingli Village, Wanrong Township, Hualien County (花蓮縣萬榮鄉明利村大加汗 (上明利) 部落), but their language usage and characteristics are not well-documented. Only few words were recorded in Tashiro (1900).

1.1.2 Central Toda

Modern Central Toda dialect is distributed in the Snuwil community (or Sakura, Gungu; Chinese: 史努櫻部落, formerly known as 春陽部落), Chunyang Village (春陽村); Toda community (都達部落 or formerly known as 平靜部落), and Ruku Daya community (平和部落) in the Toda village (都達村, formerly belonged to Chingying village (精英村)) of Ren'ai Township, Nantou County (南投縣仁愛鄉).

The current location of the Snuwil community was originally the Gungu community of the Tgdaya group. After the Musha Incident, the Tgdaya people left, and the Japanese authorities assigned this area to the Toda group. The residents of this community were originally from the Rucaw, Bngbung, Cka, Tnbarah, and Ruku Daya communities, who migrated to this location and formed the Snuwil community (Yap 2011, Ongoing).

The remaining residents of Rucaw, Bngbung, Cka, and Tnbarah merged to form the Toda community, which continues to exist today. The remaining residents of Ruku Daya stayed in the local area, and after the World War II, some of the displaced residents returned to live there (Yap 2011 and my field notes).

1.1.3 Eastern Toda

The Toda people who migrated eastward are known by different names, such as Tawsa, Tawsay, Tuda, etc. Some of them assimilated into the Klesan Atayal group (泰雅族南澳群) and switched to speak different variants of Yilan Creole (Chien and Sanada 2010; Liao 1977), whereas others gradually moved southward to their current location in Tawsa (or Tawsay) community, Lishan Village, Chuohsi Township, Hualien

County (花蓮縣卓溪鄉立山村山里部落).

The Toda group who migrated eastward, believes they originated from Tawsa-Truwan, approximately in the present-day Toda village area (南投縣仁愛鄉都達村). The migration path passed through areas such as today's Sqoyaw community of Atayal (環山部落, Seediq: Sqawraw), in Pingdeng village, Heping District, Taichung City (臺中市和平區平等里). There is a certain degree of similarity in clothing and accessories, with more diverse colors in female attire compared to other Seediq groups, including elements of purple, pink, and others.

There are some linguistic records of their language. Lee (2015) introduces its phonological system and vocabulary, she shows that the has features of Toda but it is also influenced by Truku, making it a contact-induced dialect.

1.1.4 Central Truku

Modern Central Truku dialect is used by four communities, all located in Ren'ai Township, Nantou County: Bwarung community (廬山部落) in Chingying Village (精英村); Truwan community (德魯灣部落, formerly known as 平生部落) and Sadu community (沙都部落, formerly known as 靜觀部落) in Truku Village (德鹿谷村, formerly known as Hecuo Village (合作村)); Pulan community (or Inago; 松林部落) in Chin'ai Village (親愛村).

Central Truku originally developed into five communities, including Truwan (Truku-Truwan), Sadu, Busig Ska, Busig Daya, and Brayaw. After the Musha Incident in 1930, the Japanese authorities first relocated some residents from communities other than Brayaw to where is now Pulan (Inago). Subsequently, the entire Brayaw community and some residents from other communities were moved to the Bwarung community,

originally belonging to Tgdaya. The remaining residents of Truwan, Sadu, Busig Ska, and Busig Daya merged to form the Truku community. After World War II, it further divided back into Truwan and Sadu (Yap 2011, Ongoing).

1.1.5 Eastern Truku

The Truku people who migrated to the Hualien area developed a total of 28 communities. Using a more inclusive calculation, the number can reach approximately 100, as mentioned by Liao (1978). Due to space constraints, the complete list is not provided here.

Liao (1977) suggests that this group originated from Truku-Truwan (near present-day Truwan and Sadu, located in Truku Village, Ren'ai Township, Nantou County), and they first established the Tpuqu community before starting to disperse. According to historical records and oral history, Liao (1978: 200) speculates that these communities arrived in the eastern region around or earlier than the mid-18th century.

After centuries of separation, this group of people developed a strong ethnic identity and urged independence from the Atayal Nation to form the Truku Nation in 2004. They refer to their ethnolect as "Kari Truku (太魯閣語)".

1.2 Research questions and goals

As mentioned earlier, there is currently a lack of comprehensive comparisons of Seediq dialects and reconstruction of Proto-Seediq in academic research. My recent studies highlight the necessity of reconstructing Proto-Seediq for understanding both the internal relationships of Seediq and those at a higher level (Atayalic). Additionally, there is no literature found that discusses the subgrouping and internal relationships

of Seediq dialects based on a top-down approach. Most literature either focuses on superficial dialectal differences or employs criteria such as impression and mutual intelligibility for subgrouping (relevant discussions will be presented in Section ??).

Therefore, the main research questions of this thesis are: What were the phonological and morphosyntactic systems of Proto-Seediq like? And how can the internal classification of Seediq be done?

Based on the research questions, this thesis sets three main goals to address the aforementioned research gaps and remained issues: (1) Reconstruct the phonological system of Proto-Seediq and conduct a preliminary exploration of its morphosyntactic system and (2) Investigate the subgrouping of Seediq dialects with their shared innovation(s) based on the reconstruction of Proto-Seediq.

1.3 Methodology

1.3.1 The Comparative Method — for bottom-up reconstruction and subgrouping

This paper mainly employs the Comparative Method, a fundamental tool in historical linguistics, for the purpose of reconstructing Proto-Seediq and of subgrouping Seediq dialects. The Comparative Method is a systematic approach that involves several essential steps in order to uncover linguistic relationships and reconstruct ancestral languages (Fox 1995).

The first step in this process is the collection of words from different Seediq dialects. By comparing these words, I aim to identify cognates, which are words that are similar or same in both form and meaning. The similar words among dialects also have

to show regular sound correspondences in basic vocabulary. In addition, it is necessary to exclude the possibilities of "chance", "universals", and "borrowing". Cognates provide valuable evidence for language comparison and reconstruction.

Once cognates are identified, the next step involves setting up sound correspondences. Sound correspondences are regular patterns observed across cognate words. I will group these correspondence sets together based on their phonetic similarity.

The third step employs a bottom-up reconstruction of protophonemes, which are the phonemes believed to have existed in the proto-language. In the process of reconstructing phonemes of a proto-language, careful attention must be given to the directionality of sound changes. Two fundamental reconstruction principles must be taken, as pointed out by Crowley and Bowern (2010: 85). Firstly, the sound change should be plausible and natural in language developments. Secondly, a reconstructed phoneme should minimize changes in both places and manners of articulation. By adhering to these principles, it becomes possible to reconstruct protophonemes for each corresponding group and identify consistent sound changes from the proto-language to its descendant languages.

Also, sound changes occurring from the proto-language to the descendant languages or dialects adhere to the principles of the Neogrammarian Hypothesis, alternatively recognized as the Regularity Hypothesis. According to this notion, sound changes follow a strict rule without any exceptions (Brugmann and Osthoff 1878). In essence, a sound change is anticipated to be universally applied to all forms under defined conditions in a language or dialect. Irregular cases are typically observed solely in loanwords, sporadic sound changes, and accidental possibilities. For the phoneme inventory in the proto-language, we also expect it to be a balanced system.

Using the protophonemes as a guide, I will also reconstruct the lexicon of Proto-Seediq. By applying the sound correspondences and protophonemes to the cognate words, they can infer the likely forms of words in Proto-Seediq. This step enables the reconstruction of a substantial portion of the lexicon.

Finally, the Comparative Method allows for the determination of subgrouping within the Seediq language. The subgrouping hypothesis in this paper is carried out based on the following criteria: (1) The primary method of subgrouping involves the use of exclusively shared innovations to form a subgroup, while shared retentions must not be employed as subgrouping criteria. (2) Phonological evidence includes conditioned shared sound changes, unique mergers, and sporadic sound changes. Sporadic or rare sound changes are preferred over common ones as subgrouping criteria, as common sound changes might occur independently in closely related languages, representing a phenomenon known as "drift" or "convergence among genetically related languages", such as the "Umlaut" phenomenon in West-Germanic languages. An illustrative case of this phenomenon is observed in the pluralization of English and German nouns, which involve vocal mutations, as known as "Umlaut". For instance, ***foot*** [ʊ] becomes ***feet*** [i] in English, and ***Fu**ss* [uː] turns into ***Fü**sse* [ʏ] in German. Both languages employ vowel changes to denote plurality, yet process arose independently, not a shared innovation. This evolution traces back to a common historical pattern where the stress was predominantly on the first syllable, weakening the subsequent syllable and leading to changes in the final vowels of plural nouns. Although it starts with minor sound alterations, this kind of linguistic evolution, driven by structural similarities, can result in significant changes within separate but related languages (Greenberg 1957: 47–48, Sapir 1921). (3) The lexical evidence referred to in this paper

includes categories like lexical innovations, fossilized infixation(s) within words (see Section **??**), and others.

The paper will simultaneously consider the various aforementioned types of evidence as the basis for subgrouping, as situations might arise where a combination of two or more types of evidence is needed for judgment. These subgroups represent distinct historical stages or divisions within the Seediq language.

Overall, the application of the Comparative Method in this study of Seediq dialects involves collecting data, identifying cognates, establishing sound correspondences, reconstructing protophonemes and the lexicon, and determining subgroups. By following these steps, we are able to gain valuable insights into the historical development and relationships within the Seediq language, contributing to our understanding of its linguistic evolution.

1.3.2 Top-down reconstruction method

The top-down reconstruction method is also known as "inverted reconstruction" (Hockett 1958: 512–16; Anttila 1972: 346).

For example, consider a language family shown in Figure **??**. Sometimes evidence for *Y cannot be found in language A and language B; we can use evidence from *X, which was reconstructed based on *Z and its other sister language(s) to reconstruct *Y form above (Fox 1995: 88).

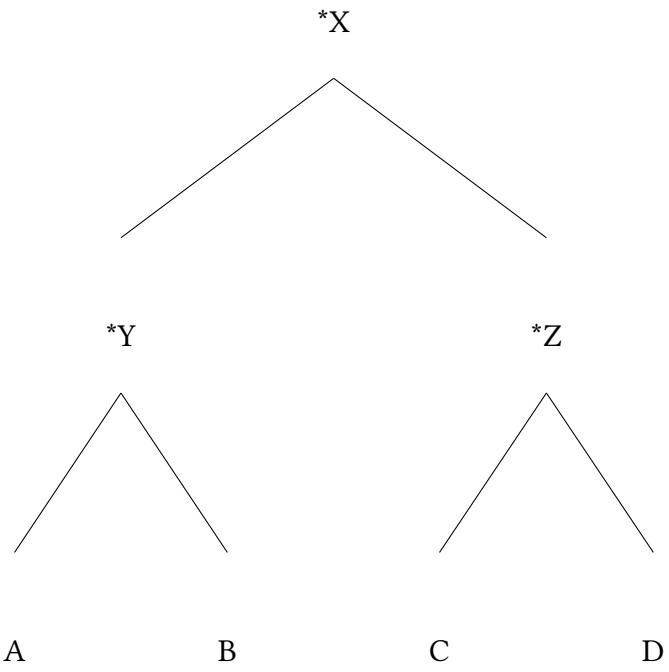


Figure 1.3: An assumed language family as an example for top-down reconstruction

In this paper, some of the reconstruction of Proto-Seediq lexical forms is primarily guided by such a method. At times, modern Seediq dialects may exhibit two or more competing forms for a given word among them. In such cases, external evidence from Proto-Austronesian can assist in determining the Proto-Seediq form.

However, considering that Atayalic languages often display numerous lexical innovations, it is frequently challenging to find evidence directly from Proto-Austronesian. Therefore, this paper employs a methodology which is similar to that of Goderich (2020: 187–88) for the reconstruction of certain forms. If Proto-Seediq has two candidates *a and *b for a specific lexical item, and if corresponding forms *b′ or b′′ exist in Proto-Atayal or Atayal dialects, I will choose *b for the reconstruction in Proto-Seediq. Relevant cases will be discussed in Section **??**.

Certainly, in cases where a particular form appears in only one dialect, lacks

agreement among the majority of dialects, or its meaning aligns with the characteristics of loanwords (such as expressing a highly specific meaning), the possibility of borrowing must be considered. Given that the Atayalic subbranch only contains two languages, the risk of circular reasoning is difficult to be entirely avoided at this stage. I can only endeavor to eliminate instances fitting the aforementioned criteria as much as possible. To address this issue more effectively, a broader comparison of various Seediq and Atayal dialects, as well as an exploration of historical contact relationships between the two distinct language branches, is necessary.

1.4 Data sources

等所有田調進行完再補 Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio

metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1.5 Orthographic conventions

I apply the orthographic conventions in the paper in order to make an easier comparison of Seediq dialects. In general, the symbol used will be the same as those in the International Phonetic Alphabet (IPA), but there are still some exceptions. The conventions are shown in Table **??**, and the current orthographic systems (MOE/CIP) that is used by the Seediq and Truku peoples are also listed for reference.[2]

The symbol *g* represents voiced velar stop or fricative [g~ɣ], as there are variations across different dialects. The symbols c and ɟ respectively denote [ts~tɕ] and [ɟ~dʑ]. In Tgdaya and Central Toda, *c* is phonemic, while *c*/*ɟ* in Eastern Toda, Central Truku, and

[2]MOE and CIP are Abbreviations for Ministry of Education and Council of Indigenous Peoples of Taiwan, respectively.

Table 1.1: Orthographic conventions

This thesis	MOE/CIP	IPA	This thesis	MOE/CIP
p	p	[p]	r	r
t	t	[t]	l	l
k	k	[k]	w	w
q	q	[q]	y	y
b	b	[b]	a	a
d	d	[d]	i	i
g	g	[g~ɣ]	u	u
c	c	[ʦ~ʨ]	e	e (Tgdaya)
ɟ	j	[ɟ~ʥ]	ey	ey (Central Toda, Central Truku, Eastern Truku)
s	s	[s]	ay	ay (Central Toda, Eastern Toda, Central Truku, Eastern
x	x	[x]	o	o (Tgdaya)
h	h	[ħ]	ow	o/ow (Central Toda, Central Truku, Eastern Truku)
m	m	[m]	aw	aw (Central Toda, Eastern Toda, Central Truku, Eastern
n	n	[n]	ə	∅/e (Central Toda, Eastern Toda, Central Truku, Eastern
ŋ	ng	[ŋ]		

Eastern Truku represents palatalized allophones of /t/ and /d/. As for *h*, it represents the voiceless pharyngeal fricative [ħ]. However, in certain (sub-)dialects or among speakers, it is also pronounced as a glottal [h]. Nevertheless, [ħ] is generally considered a more conservative pronunciation.

In the standard orthography, the velar nasal /ŋ/ is typically expressed as *ng*, but for the purposes of this thesis, it will be designated with the IPA symbol *ŋ*. The symbol *r* represents the voiced alveolar tap [ɾ]. As the sole rhotic sound in this language, this sound may have other variations, but they do not significantly impact the phoneme inventories of different dialects or the reconstruction of Proto-Seediq. *l* represents the voiced alveolar/dental lateral (or lateral fricative) [l~ɮ], and *y* denotes the voiced palatal approximant ("glide") [j].

There is an issue about the "neutralized" (or reduced) vowel in prepenultimate syllables. In general, any kinds of vowel reduced into [u] (Tgdaya)/[ə] (non-Tgdaya) in this specific environment. Since the vowel quality is predictable, it is omitted in the MOE/CIP system. For example, *dtnqlian* (Tgdaya) 'a group of slaves' is pronounced as [dutunuquli(j)an]. In this paper, the "omitted" vowel in the MOE/CIP system will be transcribed based on the phonological system of each dialect.

For the *e [ə] used in Blust's reconstruction of Proto-Austronesian (Blust et al. 2023), this thesis will directly use *ə as the corresponding symbol.

Modern languages (including Seediq and other languages) or languages with written records are marked in italics in this thesis. Reconstructed forms are marked with a single asterisk (*), and double asterisks (**) are used for the following categories: (1) forms of Pre-Proto-X, (2) intermediate stages in sound change processes, and (3) other hypothetical/expected forms that are unattested.

1.6 Outline of the thesis

This thesis is organized as follows: Chapter **??** provides a literature review on three aspects related to the Seediq language, including external relationships, internal relationships, and relevant anthropological studies. Chapter **??** introduces the phonological systems of modern Seediq dialects, encompassing phoneme inventories, phonotactics, and phonological rules. Chapter **??**, based on the synchronic phonologies presented in Chapter 3, focuses on the reconstruction of the phonological system of Proto-Seediq. Chapter **??** preliminarily reconstructs the morphosyntactic system of Proto-Seediq, including its affixes and pronominal systems. Chapter **??** discuss some issues on reconstructing the lexicon of Proto-Seediq. Chapter **??** explores the proposed subgrouping of Seediq dialects and presents relevant evidence. Chapter **??** serves as the conclusion, summarizing the thesis and discussing remaining issues and potential future research directions.

Chapter 2

Literature review

This chapter discusses previous studies related to the historical development of Seediq. Section **??** provides critiques of previous reconstructions related to Proto-Seediq and Proto-Atayalic, such as: Li (1981) and Ochiai (2016). Section **??** concentrates on reviewing relevant literature on dialectal differences and internal relationships of Seediq, like: Ogawa and Asai (1935), Cheng, Nanu Ruwiq, et al. (2017), Cheng, Dakis Pawan, et al. (2019), and Sung (2018). Section **??** discusses anthropological amd sociology research related to Seediq, seeking additional historical materials, such as migration patterns (Liao 1977, 1978; Yap 2023). Finally, Section **??** offers a summary of this chapter, including insights for the thesis.

2.1 Reconstruction

In the past, there has been almost no focus on the reconstruction of Proto-Seediq. Relevant literature mainly includes Li (1981), who reconstructed Proto-Atayalic, and Ochiai (2016), who reconstructed some Proto-Seediq lexical items while interpreting the Bullock (1874) data related to the Bu-hwan dialect.

2.1.1 Li (1981)

Li (1981) compared Atayal and Seediq dialects to reconstruct the phonology of Proto-Atayalic as well as approximately 300 lexical items. The phoneme inventory of Li's Proto-Atayalic is shown in Table **??**

Table 2.1: Li's (1981) Proto-Atayalic phoneme inventory (adapted from Li 1981: 272)

p	t		k	q		ʔ		
b	d	g'	g				i	u
		c					ə	
		s	x		h		a	
m	n		ŋ					
	l	r					aw, ay, uy	
w		y						

The paper does not delve much into Seediq itself, but Li does mention a couple of points where Seediq is considered more conservative than Atayal: (1) non-final /b/, /d/, /g/ in Seediq remain voiced stops, whereas in most Atayal dialects, they lenited into fricatives /β/ (< *b), /ɣ/ (< *g), and rhotic /r/ (< *d), (2) Seediq /r/ is more conservative (< *r), as most Atayal dialects (except for Palŋawan (Plngawan)) typically change it into other sounds (see Li 1981: 264–265 for details).

Another point worth noting is the distinction between Proto-Atayalic word-final *-aw and *-ag, which contrast in some Atayal dialects. However, in all four Seediq dialects, their correspondences are as follows: Toŋan *-o* : Toda *-aw* : Truwan *-aw* : Inago *-ag*.

Li's argument is that Inago retained Proto-Atayalic *-ag as *-ag* , while the sound

change of Proto-Atayalic *-aw to *-ag* is an innovation in this dialect (actually a merger) (Li 1981: 257–258, 271). However, this can not necessarily guarantee that the contrast between *-ag and *-aw persisted at the Proto-Seediq stage, since all dialects show a merger of the two.

Li's reconstruction marked the beginning of comparative research on Atayalic languages and dialects, providing essential reference material. However, as mentioned earlier and according to the arguments presented by Song (2022, 2023), the reconstruction of the lower level Proto-languages, remains crucial. Just as Goderich's (2020) reconstruction of Proto-Atayal has addressed some issues, the reconstruction of Proto-Seediq is the primary goal of this thesis.

2.1.2 Ochiai (2016)

The main target of Ochiai's (2016) paper is to analyze the Bu-hwan vocabulary recorded by Bullock (1874) in comparison with Paran (equivalent to Tgdaya) Seediq and Taroko (equivalent to Eastern Truku) Seediq to elucidate the historical and linguistic relationships with other Seediq dialects.

Even though the main focus of this paper was not the reconstruction of Proto-Seediq, she still reconstruct nearly 200 lexical items based on the lexical items form the three dialects.

However, the reconstructions in her paper were not done following the basic Comparative Method (see Section **??**). No sound correspondences were proposed; instead, the reconstructions were based on individual lexical items one by one. Ochiai also noted that these reconstructed forms are still tentative and some require further revision. Therefore, the appropriateness of the forms reconstructed in her paper is not

discussed further here.

An interesting point is her differentiation between *ð and *y in the reconstruction of Proto-Seediq. I have organized relevant correspondences based on her reconstructions as shown in Table **??**.

Table 2.2: Ochiai's Proto-Seediq *ð and *y

Ochiai's Proto-Seediq	Bu-hwan	Paran (Tgdaya)	Taroko (Eastern Truku)
*ð	z, dz, th	y	y
*y	i, ∅ (V_V)	y	y
	N/A	y	y

When a set of cognates has Bu-hwan *z*, *dz*, or *th* corresponding to *y* in both modern dialects, Ochiai reconstructs Proto-Seediq *ð. If Bu-hwan *i* or ∅ (V_V) corresponds to *y* in modern dialects, it is reconstructed as *y. Additionally, when there is no cognate found in Bu-hwan, it is also reconstructed as *y. This correspondence may seem plausible, but it can also lead to redundancy. Therefore, we can examine whether there is a necessity for reconstructing two protophonemes through external evidence.

I examined the words containing *z*, *dz*, *th* mentioned by Ochiai in Bullock (1874) and those containing *ð* recorded by Asai (1953) (his data was collected in 1927), comparing them with Proto-Atayal from Goderich (2020) and modern Atayal dialects. I found that out of 28 examples, only 5 were recorded as *y* (Bullock) or *j* (Asai), while the rest were *z/dz/th* (Bullock) or *ð* (Asai). Additionally, in the 21 cases where cognates in Atayal were found, Bu-hwan (1874)/Paran (1927) *z/dz/th* or *ð* corresponded to Proto-Atayal *y, *w (u_a), *g, *ʔ, *s. Examples are shown in Table **??** (the "palatal

hooks" in Asai's data are transcribed as <ʲ>).

Table 2.3: Correspondences among Bu-hwan, Paran, Proto-Atayal, and Proto-Austronesian

Bu-hwan	Paran	Proto-Atayal	Proto-Austronesian	Gloss
kasia	kasʲija	*qusiyaʔ		'water'
	bsijaq	*bVsiyaq		'long time'
	qojoqaja	*qayqayaʔ		'tool'
daia 'east'	dajja	*rayaʔ	*daya	'upslope; uphill'
măhoyesh	mɯwes	*maquwas	*quyaS	'to sing'
	nɯbɯðas	*nabuwas		'belly'
maidzah	meðaħ	*muwah		'to come'
dagizak paru	degiðaq	*ragiyax		'mountain'
sinuzuk		*sinuyug		'cord, line'
hazi kăhoni 'fruit'	heði 'fruit'	*hiʔiʔ		'flesh, meat'
ŭduthiŋ 'throat'		*wariyuŋ		'neck; nape'
	raði 'a match-maker; a relative'	sa-ɹeʔ 'to propose marriage' (Plngawan)		
	teheðaq 'to visit'	*tuwahiyaq 'far'		'far'
	hiða	*hiyaʔ	*si ia	'3SG'
	piða	*pisaʔ	*pija(x)	'how many'
swadzu makaidil 'sister'	sɯwaði	*suwaʔiʔ; *suwagiʔ 'younger sibling's spouse'	*Suaji	'younger sibling'
	ŋaðan		*ŋajan	'name'
	kiða 'that; later'	*kisaʔ 'today; soon'		'soon; later'
	dahiða	ləhəgaʔ (Squliq)		'3PL'
kuzu		*məquʔ (<**məquʔuʔ ?)		'snake'
	gaða 'customary laws'	*gagaʔ 'culture; tradition; law; religion'		'law'
dzadzuŋ	ðaðɯŋ	(*gaʔuŋ)		'river'

kuzuch	kɯðɯħ	(*quwalax)	*quzaN	‘rain’

The correspondences in Proto-Atayal may seem messy, but they can actually be divided into two groups. *y and *w (u_a) can be integrated because the Atayal *-uwa- sequence appears to have originated from an earlier stage **-uya-, for example: Proto-Austronesian *quyaS > Proto-Atayal *quwas ‘song’. On the other hand, *g/*ʔ/*s all derive from Pre-Proto-Atayal **g and represent the merger of Proto-Atayalic *g/*j (see Song 2023 for details). Related to the examples here is the reflexes of Proto-Atayalic *j.

In other words, Bu-hwan *d/dz/th* and Paran *ð* in 1927 correspond to the reflexes of Proto-Atayalic *y and *j in Proto-Atayal. Furthermore, Song (2023) proposed that Proto-Atayalic *y and *j merged in Proto-Seediq as *y, consistently reflected as *y* in all modern Seediq dialects with only few exceptions.

As mentioned earlier, in the data recorded by Asai (1953) for Paran, the majority is *ð*, which correspond to Pre-Proto-Atayal **y (< Proto-Atayalic *y) and **g (< Proto-Atayalic *j/g). Based on this, it can be concluded that Paran also underwent the Proto-Atayalic *y/*j merger, which is the same as other modern Seediq dialects.

In summary, it may not be necessary to reconstruct *y and *ð as two proto-phonemes at the Proto-Seediq level. I propose using *y [j] as the Proto-Seediq protophoneme because [j] > [ð] is a more common sound change compared to [ð] > [j]. For the reflexes of Proto-Seediq *y in modern dialects, please see Section **??**. Regarding the five examples in Bu-hwan/Paran showing *y/j* [j], since they are relatively rare, they could potentially be borrowed from neighboring dialects (at the time, the Tgdaya group had a maximum of 12 communities, as mentioned in Section **??**). However, this issue may require further investigation and research.

2.2 Internal relationships of Seediq dialects

In previous studies, only few of them discussed the subgrouping of the Seediq language. Most of the works take Atayalic as a whole, without a deeper look into Seediq. For example. Li (1981) mentions, "The Sediq dialects, Toŋan, Toda, Truwan and Inago, which are not completely uniform, but also not very divergent[.]" (p. 235), but also describes: "As there are significant dialectal differences in Sediq, I have included four dialects in this study" (p. 236).

In this section, I will discuss the primary literature that has addressed subgrouping, or simply dialectal differences in the past, including Ogawa and Asai (1935), Cheng, Dakis Pawan, et al. (2019), Cheng, Nanu Ruwiq, et al. (2017), and Sung (2018).

2.2.1 Ogawa and Asai (1935)

Ogawa and Asai (1935) is a book introduces the indigenous languages (referred to as "蕃語 ('savage's language')" at that time) during the Japanese colonial period in Taiwan, which includes 12 Austronesian languages that were spoken in Japanese colony. Its content includes overviews, grammatical sketches of each language, as well as myths and stories.

Ogawa and Asai (1935) point out that Seediq is the language of the Seediq ethnic group that has been living in *Taichūshū* (臺中州) and *Karenkōchō* (花蓮港廳) for a long time.

They argue that Seediq can be roughly divided into two dialects: 霧社方言 (Musha dialect, belongs to 霧社蕃 (Musha)) and タロコ方言 (Taroko dialect, including トロコ蕃 (*Toroko*, (Central) Truku), タウツア蕃 (*Tautsa*, Tawca), タウサ

イ 蕃 (*Tausai*, Tawsay), タ ロ コ 蕃 (*Taroko*, (Eastern Truku)), バ ト ラ ン 蕃 (*Batoran*, Btulan)). Table **??** shows the phonological differences between two dialects (including their subdialects).

Table 2.4: Phonological relationship between Musha and Taroko dialects in Ogawa and Asai (1935)

Musha		Taroko
Paran	Gungu	Ibuh
ð	y [j]	y [j]
ts (non-final)	ts (non-final)	s (non-final)
g?	q/w	g [ɣ]
g	g	g

Based on their description, the subgrouping tree can be drawn as is Figure **??** (the spelling of each group/village is transcribed based on modern orthography). However, Ogawa and Asai did not explain the criteria they used for classification; they only provided simple descriptions of the differences among dialects.

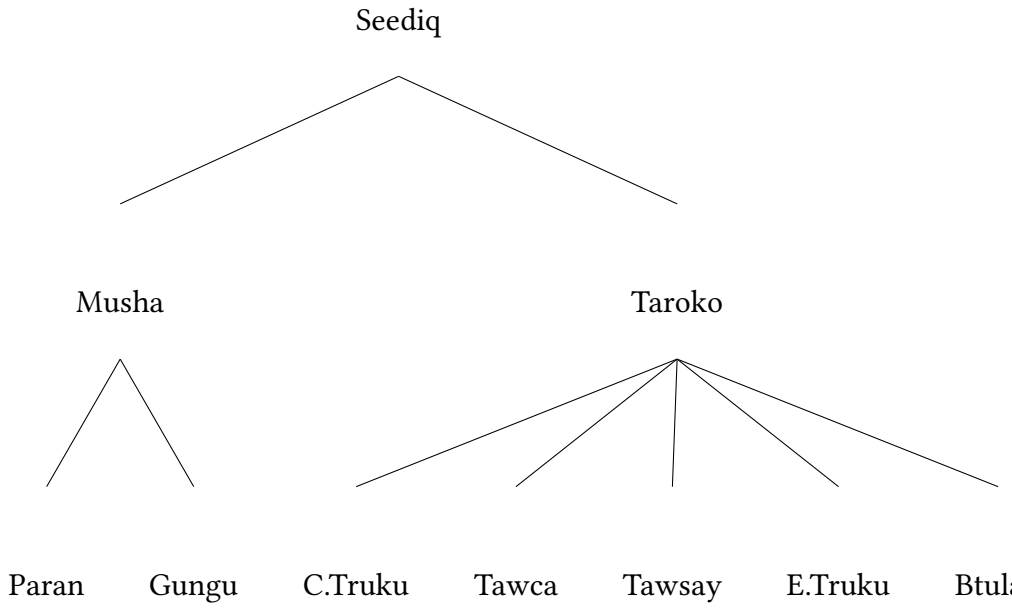


Figure 2.1: Seediq subgrouping tree based on Ogawa and Asai's (1935) description

2.2.2 Cheng, Nanu Ruwiq, et al. (2017) and Cheng, Dakis Pawan, et al. (2019)

Cheng, Nanu Ruwiq, et al. (2017) and Cheng, Dakis Pawan, et al. (2019) respectively refer to the vocabulary handbooks of (Eastern) Truku and Tgdaya. As both of them share exactly the same subgrouping hypothesis, they are discussed together in this context.

Cheng, Nanu Ruwiq, et al. (2017) and Cheng, Dakis Pawan, et al. (2019) mention that Seediq has three ethnic subgroups: Tgdaya, Toda, and Truku. However, they indicate that Seediq can be roughly divided into two linguistic dialects: Tgdaya (德固達雅方言) and Truku-Toda (太魯閣/都達方言). Truku-Toda can be further divided into 太魯閣土語 (Eastern Truku), 德路固土語 (Central Truku), 都達土語 (Toda), and 見晴土語 (Miharasi). See Figure **??** for their subgrouping tree.

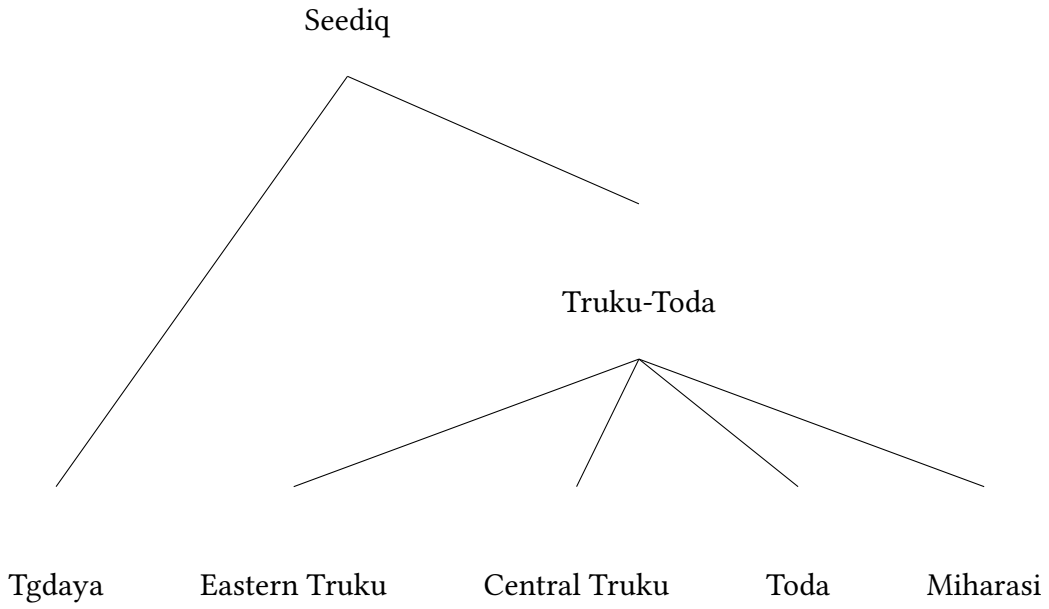


Figure 2.2: Seediq subgrouping tree based on description of Cheng, Nanu Ruwiq, et al. (2017) and Cheng, Dakis Pawan, et al. (2019)

Eastern Truku in Hualien and Central Truku in Nantou are slightly different due to the geographic isolation, but they can still communicate with each other without problems. Toda is different from Truku, but not a lot; they has mutual mutual intelligibility. There are also differences between Miharasi and Eastern Truku. (Cheng, Nanu Ruwiq, et al. 2017: 4; Cheng, Dakis Pawan, et al. 2019: 5)

In Cheng, Nanu Ruwiq, et al. (2017) and Cheng, Dakis Pawan, et al. (2019), there is no evidence presented for subgrouping of Seediq dialects. Personally, I speculate that they might have used mutual intelligibility as a criterion for the subgrouping (or the "division of dialects"). However, the main issue with using this criterion for subgrouping is that the causes of mutual intelligibility could be due to shared retentions, convergence resulting from language contact, and so forth, which are not the best indicators for subgrouping. As discussed in Section **??**, shared innovations are the pre-

ferred criterion for subgrouping, not shared retentions. Therefore, the model provided by them needs to be revised.

2.2.3 Sung (2018)

Sung (2018) is a reference grammar book for Tgdaya Seediq, providing clear descriptions of Seediq grammar. However, only certain paragraphs mention phonological differences among dialects, summarized in Table **??**.

Table 2.5: Seediq dialectal differences (based on Sung 2018: 21–22)

Tgdaya	Toda	Truku	Note
e	ə	ə	penultimate
-e	-ay	-ay	final
-o	-aw	-aw	final
t, d	t, d	c, ɟ	/_i
g-/-g-	w-/-w-	g-/-g-	non-final
-w	-g	-g	final
c	c	s	non-final
c	c	c (Central Truku) / t (Eastern Truku)	final

Sung (2018: 22) mention that the occurrence of -g at the word-final position in Toda. This is likely a typo, as there is no /g/ in the native phoneme inventory of Toda (it only appears in loanwords).

Other than the aforementioned point, this section in the book provides the main differences among Seediq dialects. However, without reconstructing the Proto-Seediq

phonology, we cannot determine whether these dialects can form any subgroup, and this is not the focus of Sung's book, so I only examine her description of dialectal differences.

2.3 Relevant anthropological and sociology studies

Understanding the history of ethnic groups is also crucial for understanding the changes and development of languages. This section will examine past research in anthropology, sociology, and related fields on oral history and migration (whether voluntary or forced), to explore the relationships between groups or the background of language contact.

The primary references include Utsurikawa et al. (1935), Liao (1977, 1978), and Yap (2023). Most of the history and migration of Seediq(-Truku) groups have been extensively covered in Chapter **??**. Therefore, this section will focus on reflecting on the aspects where oral history and migration history may have an impact on language.

2.3.1 Utsurikawa et al. (1935)

Utsurikawa et al. (1935) is a book that documents the history of various indigenous nations through field works and interviews to study their respective systems. At that time, Seediq was classified as a branch of Atayal, but the authors also conducted interviews with each Seediq group. However, due to the outbreak of the Musha Incident in 1930, information about the Seediq is quite limited.

According to the historical accounts transmitted by the Paran community, Paran, Suku, Gungu, Mhebu, Tongan, Sipo, Tkanan, and Qacuq successively migrated from the oldest Tgdaya Truwan community. Additionally, Drodux branched off from Gungu.

Furthermore, two other communities belonging to the Tgdaya group, Bwarung and Bkasan, originally part of the Toda group, joined the Tgdaya group around the 1900s (Teng and Jian 2023: 146–157), but there are currently no records found regarding the languages of these two communities.

The Truku group migrated upstream along the Jhuoshuei River (濁水溪, Seediq: *Yayung Mtudu*) from low-altitude areas, eventually settling in the current location of Truku Truwan. Later, the Sadu community branched off from Truku, and from Sadu, two more communities, Busig Ska and Busig Daya, emerged. Subsequently, Brayaw community split off from Busig Daya. When Sadu separated from Truku Truwan initially, part of the population migrated eastward, which is the origin of today's Truku Nation (太魯閣族).

The Toda (Tawda) group has ancestral legends suggesting an initial division between the Bkasan and Wahar communities. Eventually, most of the people from Wahar were persuaded to settle in Bkasan, while another group migrated eastward, becoming the Tawsay. Another group from Wahar migrated all the way to Tnbarah, where they later split into five communities: Cka, Bngbung, Rucaw, Ruku Daya, and the original Tnbarah. Regarding the eastward migration of the Tawsay group, some accounts suggest they initially traveled north through the territory of Sqawraw (Atayal: Sqoyaw) before continuing their journey eastward.

2.3.2 Liao (1977, 1978)

Liao (1977, 1978) delve into the ancestry and origins of the migrating Seediq people. Overall, he provide much more detailed information compared to Utsurikawa et al. (1935).

Liao (1977) also mentioned that the migrating Tawsay group (Eastern Toda) passed through the territories or hunting grounds of Slamaw and Sqoyaw during their migration process before continuing eastward to the present-day Hualien area in the north. Subsequently, they were invaded by the eastward-migrating Truku people, leading them to reside with the Klesan Atayal Group in the north. According to Liao (1977: 65), elderly individuals aged 50-60 (estimated to be born around the 1920s) could still speak their language, but younger generations had lost the ability. Around 1930, the Japanese authorities forcibly relocated the remaining Tawsay group members to live with the Truku people. The last group of Tawsay people migrated southward to the present-day Tawsa community location (see Section **??**), which is the last surviving community mainly composed of the Tawsay people (and their language, see Lee 2015).

The eastward-migrating Tgdaya group (referred to as Pribaw) settled close to the territory of the Amis villages, serving as intermediaries for trade between the eastern and western regions. Specifically, the Eastern Tgdaya people purchased goods produced by the Amis and brought them back to Puli (埔里) for trading what Amis people want to buy, until the Eastern Truku grew stronger and occupied their territory (Liao 1977: 70).

Some Truku groups migrated eastward through the saddle between the eastern peak of Mount Hehuan (合歡山東峰) and the northern peak of Mount Chilai (奇萊主山北峰), establishing the first Tpuqu community. They then rapidly expanded, creating dozens of subsidiary communities, forming the strongest branch of Seediq in the east and compelling the Tawsay and Pribaw to migrate elsewhere. Their migration path was not far from the territory of the Atayal tribe to the west, suggesting possible historical interactions.

2.3.3 Yap (2023)

Yap (2023), leveraging his expertise in sociology and geography, explores the fundamental changes in the "group resettlement policy (集團移住政策 *shūdanijūseisaku*)" within Japan's so-called "savage-managing policy (理蕃政策 *ribanseisaku*)" before and after the 1930 Musha Incident. What sets Yap's work apart from previous ones is that he does not simply investigate migration in a linear manner. Instead, he delves into the relational study of network changes between communities, examining the impact of group resettlement policies on such network relationships.

Yap (2023: 137, 153) mentions that the Tausa (Tawsay) group and the Truku (Eastern) group were enemies and were in a disadvantaged position. Their connection with the central Toda group was also severed by the Truku people. As a result, they turned to establish friendly relations with the Atayal (Klesan group) in Nan'ao, Yilan, forming a collective group in terms of geopolitics. However, later on, the three Tausa communities were relocated southward to the present-day Tawsa community (see Section **??**), interrupting their close contact with the Atayal. Nevertheless, we can still discern from this that the relationship between the Eastern Toda (Tawsay, Tausa) and the Atayal was once quite close.

2.4 Interim summary

Based on the discussion in this chapter, we have conducted a preliminary review of the historical development of Seediq. In terms of language, we have examined Li's (1981) Proto-Atayalic and Ochiai's (2016) Proto-Seediq reconstruction, discussing research gaps and parts that warrant revision.

Regarding dialect relations, a basic understanding has been established through a review of literature from the past century. At the same time, it has been observed that there is currently no subgrouping hypothesis proposed based on solid linguistic evidence.

Literature from other relevant disciplines provides insights into internal community relations and potential external contacts of the Seediq, and this thesis will also reference relevant historical backgrounds to explore the internal and external relationships of the Seediq language.

Chapter 3

Phonological sketches of Seediq dialects

This chapter provides the synchronic phonological sketches of five Seediq dialects: Tgdaya, Central Toda, Eastern Toda, Central Truku, and Eastern Truku. Section **??** discusses the phoneme inventories and phonotactics of all the five dialects. In section **??**, different kinds of alternations are discussed, especially those related to morphophonology.

3.1 Phoneme inventories and phonotactics

This section introduces the phoneme inventories and phonotactics of various Seediq dialects. While the consonant inventories of Seediq dialects are similar, there are still differences; however, the vowels, particularly in Tgdaya, exhibit more distinctiveness.

Regarding phonotactics, there is a high level of consistency among dialects in terms of constraints on the occurrence of some sounds in certain positions. Stress placement is non-phonemic, with the majority of cases regularly falling on the penul-

timate syllable, while Central Toda exhibits a stress shift when the penultimate vowel is schwa /ə/ (See Section **??**).

3.1.1 Tgdaya phonology

3.1.1.1 Tgdaya consonant inventory

Tgdaya has the largest number of consonants among all dialects. The Tgdaya dialect has eighteen consonants, including seven stops *p, t, k, q, b, d, g*, one affricate *c* [ʦ], three fricatives *s, x, h* [ħ], three nasals *m, n, ŋ*, one lateral *l*, one tap *r* [ɾ], and two glides *w, y* [j]. Table **??** shows the consonant inventory of Tgdaya.

Table 3.1: Tgdaya consonant inventory

	Bilabial	Dental/Alveolar	Palatal	Velar	Uvular	Pharyngeal
Stop	p b	t d		k g	q	
Fricative		s		x		h [ħ]
Affricate		c [ʦ]				
Nasal	m	n		ŋ		
Lateral		l				
Tap		r [ɾ]				
Glide	w		y [j]			

<h> is realized as a pharyngeal [ħ] by more conservative speakers; on the contrary, other speakers pronounce it as a glottal [h] (also by some elder speakers). Further investigation is needed to determine whether this is influenced by age or other factors (such as societal ones). Additionally, careful phonetic studies are required. But phono-

logically, the realizations of this phoneme do not change the categorical properties, and it is clearly distinct from velar fricative /x/.

<c> /ʦ/ and <s> always palatalize to [ʨ] and [ɕ] before the high front vowel /i/, and whether it occurs in other environments depends on the speaker. Common environments include: (1) *i*_#, and (2) _*e*. For example, *madis* [madiɕ] 'to bring (AV)', *cebu* [ʨebu] 'to shoot'.

<l> is a lateral sound, but in the realization of some speakers, it may exhibit slight frication. However, due to the lack of noise, it does not reach the level of [ɮ].

<r> is usually pronounced as an alveolar (or retracted) tap [ɾ~ɾ̠] or a retroflex flap [ɽ]. However, as the only rhotic sound, there may be even more variations, such as approximants [ɹ~ɻ] or even (rarely) a trill [r].

Regarding the glottal stop [ʔ], Yang (1976) argues that it lacks phonemic status and that the phonetic realization of [ʔ] and ∅ can be in free variation. Conversely, Holmer (1996) classifies the glottal stop as the phoneme /ʔ/, but without further explanation. Sung (2018) does not consider it as a phoneme and also does not provide an explanation. Since the occurrence of the glottal stop [ʔ] can be predicted and there is no phonemic distinction from ∅, I adopt Yang's view and do not consider the glottal stop as a phoneme in this thesis. This also applies to other dialects, and I will not elaborate further on this point in this thesis.

3.1.1.2 Tgdaya vowel inventory

Tgdaya has five monophthongs: *a, i, e, u, o*, as shown in Table **??**. In addition to monophthongs, Tgdaya has a diphthong *-uy* [uj].

In Tgdaya, vowel length is not distinguished. If a sequence of two identical vowels

Table 3.2: Tgdaya vowel inventory

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

is seen, the two vowels actually belong to two separate syllables, as in *haan* [ħa.an] 'to go (LV)'.

3.1.1.3 Tgdaya phonotactics

Table **??** displays the syllable types of Tgdaya, listing the syllable types covered in slower and deliberate speech. In everyday and faster speech, other types may occur, which will be briefly introduced later on. Please note that the symbol C in the coda position includes all consonants except for glides, while G represents the two glides *w* and *y*. In the onset position, there is no distinction, and all non-vowel sounds are unified under C. This convention applies to other parts of the thesis as well.

Table 3.3: Syllable types in Tgdaya

Type	Prepenultimate	Penultimate	Final	
V	—	**i**.ma 'who'	sa.**i**	'to go.IMP (PV)'
VC	—	—	sa.**un**	'to go (PV)'
VG	—	—	(not found)	
CV	**qu**.si.ya 'water'	qu.**si**.ya 'water'	qu.si.**ya**	'water'
CVC	—	—	ba.**raq**	'lungs'
CVG	—	—	lu.**buy**	'bag'

CV is the only syllable type allowed to appear in the prepenultimate position. Onsetless V will be deleted by a universal phonological rule, as seen in *eyah* /ejaħ/ 'to come' and *yah-an* /ejaħ/ + /an/ 'to come (LV)'. In the penultimate position, only V and CV structures are allowed to appear.

In the final position, Tgdaya can at least have five structures: V, VC, CV, CVC, and CVG, with fewer restrictions. While the VC structure is speculated to be possible, examples may be relatively rare and harder to find. In other dialects, there is *dəmuuy* 'to hold in fist (AV)', but irregularly corresponds to *dumoi* in Tgdaya.

In faster speech, two other syllable structures that are commonly found and are both related to nasals can occur: CVN and N̩. When the nasal consonant with an antepenultimate vowel is followed by a homorganic consonant, syncope occurs between the nasal and the following consonant. If there is a preceding syllable before the nasal consonant, the nasal consonant becomes the coda of the preceding syllable. If there is no preceding syllable, the nasal sound becomes syllabic. The occurrences of these syllable structures can be considered surface phenomena. For example, /tunusapaħ/ → [**tun**.sa.paħ] 'family', and /mupitu/ → [**m̩**.pi.tu] 'seven'. These variations may differ depending on individuals or subdialects.

Some sounds in Tgdaya also have positional restrictions, some of them are categorical. For instance, voiced stops /b/, /d/, /g/ do not occur in word-final position. Labial consonants /p/ (including /b/) and /m/ are also not found in such position.

The phoneme /x/ only appears at the word-initial position in the word ***x**iluy* 'iron', with around ten more examples found word-medially, but it is more common in word-final position.

In a significant portion of speakers, both liquids /l/ and <r> /ɾ/ merge with /n/

in word-final position, respectively. More conservative speakers maintain the distinctions, but generally, the merger of /l/ and /n/ in word-final position is more common than that of <r> /ɾ/ and /n/. Speakers of Tgdaya who retain the word-final <r> often realize it as an approximant [ɹ~ɻ], unlike the tap/flap in word-initial and word-medial positions. Of course, there may be more variants because, as mentioned earlier, this is the only rhotic sound in the language.

The word-final <y> /j/ can only occur with /u/, as in *bab**uy*** 'pig', while in other positions, there are no restrictions. As for /w/, it is totally banned in word-final position. However, according to Yang (1976), there is a distinction between *-uw* and *-u*. The absence of this distinction can be due to an ongoing merger and generational differences.

The mid vowels /e/ and /o/ primarily serve as the nucleus of open syllables due to historical developments (see Section **??**). However, with language development or analogical changes, there may be extremely rare exceptions, such as the occurrence of *muuyas* ~ *muu**wes*** /muujas/ 'to sing; to study (AV)' in casual speech.

In Tgdaya, there is a rule of neutralization of prepenultimate vowels, with the neutral vowel being /u/. This means that, in general, vowel distinctions only occur in penultimate and final syllables, while the vowels in prepenultimate syllables are neutralized to [u]. However, when the onset of the penultimate syllable is <h> /ħ/ or has no onset (i.e., the onset is ∅), the neutralized vowel of the antepenultimate syllable is assimilated to be the same as the penultimate syllable, as in *m**e**h**e**diq* [m**e**.ħ**e**.diq] 'to give way (AV)' and *s**ii**mah* [s**i**.**i**.maħ] 'to drink (CV)'. However, there are exceptions, such as *m**u**h**u**hediq* [m**u**.ħ**u**.ħe.diq] 'to give way to each other' instead of *[m**e**.ħ**e**.ħe.diq], possibly indicating that vowel features tend to spread only once. Further synchronic

phonological analysis is needed to explore this phenomenon.

3.1.2 Central Toda phonology

3.1.2.1 Central Toda consonant inventory

The Central Toda dialect has seventeen consonant, including six stops *p, t, k, q, b, d*, one affricate *c* [ʦ], three fricatives *s, x, h* [ħ], three nasals *m, n, ŋ*, one lateral *l*, one tap *r* [ɾ], and two glides *w, y* [j]; the inventory is as given in Table **??**.

Table 3.4: Central Toda consonant inventory

	Bilabial	Dental/Alveolar	Palatal	Velar	Uvular	Pharyngeal
Stop	p b	t d		k	q	
Fricative		s		x		h [ħ]
Affricate		c [ʦ]				
Nasal	m	n		ŋ		
Lateral		l				
Tap		r [ɾ]				
Glide	w		y [j]			

Central Toda lacks the consonant <g> (/g/ or /ɣ/). Apart from this, the consonant inventory is the same as the one in Tgdaya.

<s> and <c> /ʦ/ also palatalize to [ɕ] and [ʨ] before the high front vowel /i/, but do not occur in other environments. The fricative <h> can be either pharyngeal [ħ] or glottal [h].

The lateral <l> also exhibits slight frication, similar to Tgdaya. The rhotic <r>

can also be realized as an alveolar (or retracted) tap [ɾ~ɾ̠] or a retroflex flap [ɽ], and in word-final position, it is also a tap/flap with a release, not an approximant.

3.1.2.2 Central Toda vowel inventory

Central Toda has four monophthongs: *a, i, u, ə*, as shown in Table **??**.

Table 3.5: Central Toda vowel inventory

	Front	Central	Back
High	i		u
Mid		ə	
Low		a	

Apart from monophthongs, Central Toda has three phonemic diphthongs: /aw/, /aj/, and /uj/. The diphthong /aw/ has two allophones: [ow] and [aw]. [ow] appears in word-medial positions, as in *dowriq* 'eye', while [aw] occurs in word-final positions, as in *babaw* 'above; on top of'. The two allophones are in complementary distribution. The same situation also occurs with <ey> [ej] and <ay> [aj], where the phoneme /aj/ is the allophone in word-medial and word-final positions, as in *meytaq* 'to stab (AV)' and *balay* 'true; correct'. However, <uy> /uj/ can only occur in word-final position.

3.1.2.3 Central Toda phonotactics

The syllable types of Central Toda are shown in Table **??**.

Central Toda stands out in that, in the prepenultimate position, besides CV, it can also accept the presence of a single V. However, the only vowel that can appear in this position is /u/. The cause of this phenomenon stems from a historical change, detailed

Table 3.6: Syllable types in Central Toda

Type	Prepenultimate		Penultimate		Final	
V	**u**.ci.luŋ	‘lake; sea’	**i**.ma	‘who’	sa.**i**	‘to go.IMP (PV)’
VC	—		—		sa.**un**	‘to go (PV)’
VG	—		**ow**.raw	‘k.o. bamboo’	də.mu.**uy**	‘to hold (AV)’
CV	**qə**.si.ya	‘water’	qə.**si**.ya	‘water’	qə.si.**ya**	‘water’
CVC	—		—		ba.**raq**	‘lungs’
CVG	—		**dow**.riq	‘eye’	lu.**buy**	‘bag’

in Section **??**.

As for the penultimate position, it also allows the appearance of types such as VG, CVG, as in ***ow.raw*** ‘k.o. bamboo’ and ***dow.riq*** ‘eye’. The final syllable permits all six types to occur.

The syncope phenomenon of antepenultimate vowels after a nasal mentioned earlier is also quite common in Central Toda, occurring more extensively and possibly even being obligatory for some speakers (requiring further extensive investigation). This phenomenon occurs not only when a nasal is followed by a homorganic consonant in the next syllable but also in other situations, but it seems to be limited to situations involving the infix <ən> ‘PST; NMLZ’. For example: /qənəpaħan/ → [**qəm**.pa.ħan] ‘field (lit. the place for working)’.

In the Central Toda dialect, the restrictions on the occurrence of certain phonemes are very similar to what we see in Tgdaya. For example, voiced stops /b/, /d/, labial consonants /p/ (/b/), /m/ are not allowed in word-final positions. The position and frequency of /x/ appearing in the Central Toda dialect are almost identical to those in Tgdaya. /x/ in word-initial position is only observed in the word *xiluy* ‘iron’.

The neutralization of prepenultimate vowels can also be observed in Central Toda; however, in this dialect, the neutralized vowel is schwa /ə/. Therefore, any vowel occurring before the penultimate syllable, regardless of its underlying form (or diachronically speaking, its historical source), uniformly changes to /ə/. However, there is one unique exception, where the vowel /u/ occasionally appears in that position. For example: *t**u**daya* 'the north side'. This phenomenon is related to a series of historical changes, similar to the case of prepenultimate vowels being /u/, for the same reason (see the discussion in Section **??**).

In Central Toda, there is another notable restriction that we do not see a VV (or əV) sequence across the antepenultimate and penultimate syllables. If the penultimate syllable lacks an onset (i.e., it starts with a vowel), the neutralized vowel in the antepenultimate position does not appear on the surface. For example, /səədiq/ → [sə.diq] 'human; person', /usa/ + -/ani/ → //usaani// → [sa.ni] 'to go.IMP (CV)'. This phenomenon can be explained in several ways: (1) deletion of the antepenultimate schwa /ə/ in this environment, (2) vowel coalescence after assimilation of the antepenultimate schwa /ə/ with the following vowel. Currently, only a few exceptions have been found, such as *ciida* 'at that time'.

As mentioned earlier, Central Toda has a stress shift rule. Generally, the stress falls on the penultimate syllable, but in Central Toda, if the vowel in the penultimate syllable is schwa /ə/, the stress shifts to the final syllable. However, during my fieldwork, I found that sometimes the stress still falls on the penultimate syllable in the same word, so this rule of stress shift should be considered more of a tendency rather than mandatory.

3.1.3 Eastern Toda phonology

3.1.3.1 Eastern Toda consonant inventory

The Eastern Toda dialect has sixteen consonant, including six stops *p, t, k, q, b, d*, three fricatives *s, x, h* [ħ], three nasals *m, n, ŋ*, one lateral *l*, one tap *r* [ɾ], and two glides *w, y* [j] (Lee 2015); the inventory is as given in Table **??**.[1]

Table 3.7: Eastern Toda consonant inventory

	Bilabial	Dental/Alveolar	Palatal	Velar	Uvular	Pharyngeal
Stop	p b	t d		k	q	
Fricative		s		x		h [ħ]
Nasal	m	n		ŋ		
Lateral		l [ɮ]				
Tap		r [ɾ]				
Glide	w		y [j]			

The consonant inventory of Eastern Toda is the smallest among all Seediq dialects. Besides its own mergers, some further mergers are influenced by Eastern Toda. For example, Eastern Toda does not have a phonemic <c> /ʦ/. Further explanation will be provided in Section **??**.

Eastern Toda still has two allophonic symbols: <c> [ʨ] and <ɟ> [ɟ~ʤ]. These represent the occurrence of /t/ and /d/ as allophones before the high front vowel /i/ (or homorganic glide <y> /j/), such as *cimu* [ʨimu] 'salt' and *səɟiq* [səɟiq] 'person'. List-

[1]The glottal stop /ʔ/ in Eastern Toda is treated as a phoneme in Lee (2015), but it is omitted in this thesis based on the reason mentioned earlier in Section **??**.

ing these two allophonic sounds with different symbols aims to illustrate the dialectal differences while also adhering to the conventions of past research.

The lateral <l> is a lateral fricative [ɮ] (Lee 2012, 2015). Apart from the above-mentioned differences, all other consonants in Eastern Toda are the same as those in Central Toda.

3.1.3.2 Eastern Toda vowel inventory

Eastern Toda has four monophthongs: *a, i, u, ə* (Lee 2012, 2015), as shown in Table ??.

Table 3.8: Eastern Toda vowel inventory

	Front	Central	Back
High	i		u
Mid		ə	
Low		a	

Eastern Toda has three phonemic diphthongs: /aw/, spelled <ay> /aj/, and <uj> /uj/. The most significant difference from Central Toda (and the Truku dialects soon to be mentioned) is that /aw/ and <ay> /aj/ in Eastern Toda do not undergo raising in word-medial positions (as *ow* and *ey* [ej]). In other words, whether in word-medial or word-final positions, they are realized as *aw* and *ay* [aj], as in ***aw**riq* 'eye', *sən**aw*** 'man; male', ***ay**taq* 'to stab (AV)', and *har**ay*** 'be not' (data obtained from Lee 2015).

3.1.3.3 Eastern Toda phonotactics

等田調補完剩下內容

The stress shift observed in Central Toda (when /ə/ appears in the penultimate syllable) is not found in Eastern Toda (Lee 2015: 96–97). Eastern Toda retains the stress on the penultimate syllable when /ə/ appears in that position, without shifting it to the final syllable. Lee (2015: 97) suggests that the stress shift in Central Toda is a later development, possibly to avoid stress falling on schwa /ə/. Such a phenomenon can be also found in many languages.

3.1.4 Central and Eastern Truku phonologies

Both the Central one and the Eastern Truku dialects have the same phoneme inventory. Thus, they will be discussed together here.

3.1.4.1 Central and Eastern Truku consonant inventory

The two Truku dialects have 17 consonants, including six stops *p, t, k, q, b, d*, four fricatives *s, x, g* [ɣ], *h* [ħ], three nasals *m, n, ŋ*, one lateral fricative *l* [ɮ], one tap *r* [ɾ], and two glides *w, y* [j]. Table **??** shows the inventory of consonants.

Truku lacks the phonemic affricate *c*, but the phonetic *c* can occur as an allophone of /t/ in word-final position. Table **??** only displays the consonant phonemes of Central Truku and Eastern Truku. However, there are two allophones that are distinguished in writing: <c> [ʦ~ʨ] and <ƚ> [ʤ~ʥ]. In Central Truku, <c> at word-final position phonetically appears as [ʦ], functioning as an allophone of /t/; whereas <c> and <ƚ> occur before the high front vowel /i/ in other word-initial and word-medial positions in Central Truku and Eastern Truku, respectively, serving as palatalized/affricated allophones of /t/ and /d/, as seen in examples like *cimu* [ʨimu] ‘salt’ and *səəƚiq* [səəʤiq] ‘person’.

Table 3.9: Central and Eastern Truku consonant inventory

	Bilabial	Dental/Alveolar	Palatal	Velar	Uvular	Pharyngeal
Stop	p b	t d		k	q	
Affricate						
Fricative		s		x g [ɣ]		h [ħ]
Nasal	m	n		ŋ		
Lateral fric.		l [ɮ]				
Tap		r [ɾ]				
Glide	w		y [j]			

<g> is pronounced as an fricative [ɣ] in Truku, such as *gaya* ['ɣaja] 'law, tradition', *gaga* ['ɣaɣa] 'that'.

The phoneme <l> is a lateral fricative [ɮ], a some Eastern Truku speakers pronounced it as a devoiced [ɬ] (Lee 2018: 16). In word-final position, it may merge with [n] (Lee 2018: 16, 23).

The phoneme <r> has more variants. The Central Truku speakers tend to pronounce it as a tap [ɾ], [ɾ̞], or an approximant [ɹ], while some of the Eastern Truku speakers pronounce it more like a lateral [l] (e.g. Cheng, Nanu Ruwiq, et al. (2017) transcribe this phoneme as [l], and Ogawa and Asai (1935) as *l*). Based on my personal observation, the <r> sound in Eastern Truku does indeed sound relatively longer than other dialects in duration. However, whether it carries the [+lateral] feature phonologically requires further consideration. This is merely a phonetic realization and does not currently affect the overall phonological system; therefore, I still transcribe it as <r>.

The other consonants are the same to the ones in other dialects.

3.1.4.2 Central and Eastern Truku vowel inventory

Like, Central Toda and Eastern Toda, the two Truku dialects also have four monophthongs: *a, i, u, ə*, as shown in Table **??**.

Table 3.10: Truku vowel inventory

	Front	Central	Back
High	i		u
Mid		ə	
Low		a	

Lee (2018: 12) suggests that the phoneme /u/ also has an allophone [o]. This pure [o] sound (not diphthong [ow]) is found in a very limited number of words, mostly function words, such as the topic marker *o* and the conjunction *do*. The former corresponds to *u* in Toda dialects, with a possible sporadic change in Truku dialects; while the latter may be a contraction of *da* 'CSM' and *u~o* for 'TOP'. Therefore, I currently consider it as a marginal phoneme and temporarily exclude it from the table.

Central Truku and Eastern Truku also have three phonemic diphthongs: /aw/, /aj/, and /uj/. The allophones of /aw/ and /aj/ word-medially are [ow] and <ey> [ej], respectively. Meanwhile, /uj/ also does not occur in word-medial positions. These distributions are the same as those of Central Toda.

3.1.4.3 Central and Eastern Truku phonotactics

Central Truku and Eastern Truku allow VC (which is actually VN) structures to occur in prepenultimate positions, while nasal codas must be homorganic with the

following vowel. The other syllable structure types are the same as those in Central Toda. However, VG structures in the penultimate position are only found in Central Truku, with ***ow*** in ***ow****raw* 'k.o. bamboo' possibly being the only example representing this structure, while no corresponding form is found in Eastern Truku. The syllable structures of Central Truku and Eastern Truku are shown in Table **??**.

Table 3.11: Syllable types in Central and Eastern Truku

Type	Prepenultimate		Penultimate		Final	
V	**ə**.u.da	'thing'	**i**.ma	'who'	sa.**i**	'to go.IMP (PV)'
VC	**əm**.pi.tu	'seven'	—		sa.**un**	'to go (PV)'
VG	—		**ow**.raw (CTR)	'k.o. bamboo'	də.mu.**uy**	'to hold (AV)'
CV	**qə**.si.ya	'water'	qə.**si**.ya	'water'	qə.si.**ya**	'water'
CVC	—		—		ba.**raq**	'lungs'
CVG	—		**dow**.riq	'eye'	lu.**buy**	'bag'

Some subdialects of Eastern Truku may exhibit further syncope, and the pattern is not observed in other dialects, such as /mətəɣəsa/ → [mə.təɣ.sa] 'to teach (AV)'. However, there is currently no comprehensive synchronic phonological research on this issue. This thesis also excludes these innovative features of individual subdialects from the table because they are not present in the majority or all of the Truku dialects.

As for the phonological constraints in Central Truku and Eastern Truku, they are generally similar to those mentioned in several other dialects, including the prohibition of voice stops and labial sounds in the word-final position. However, <g> [ɣ] is allowed to appear in word-final positions, but only when combined with the high vowels /i/

and /u/, as in *barig* ‘to buy’ and *elug* ‘road; path’. However, among relatively younger speakers, the endings *-ig* and *-ug* have respectively shifted to *-iy* ([-ij] or [-iː]) and *-uw* ([-uw] or [-uː]). This phenomenon is also observed in most Atayal dialects (Goderich 2020; Li 1982), and there are varying degrees of changes in other Seediq dialects (see Section **??**).

Vowels in prepenultimate positions still neutralize to schwa /ə/. They assimilate to homorganic [i] before the glide /y/ (less often to [u] before /w/). However, there are still a few exceptions, especially in the numerals ‘six’ *m**a**taru* and ‘eight’ *məsəpac* ~ *m**a**səpac* (CTR), *m**a**səpat* (ETR). These two numerals may be borrowed from Atayal, as discussed in Section **??**.

3.2 Synchronic morphophonological alternations

In each Seediq dialect, there are some phonological alternations, often resulting from the interaction between position of the sounds and morphology. Interestingly, the patterns of alternations among dialects exhibit a considerable degree of consistency. Most alternating patterns can be traced back to Proto-Seediq, implying that synchronic alternations can be viewed as remnants of the Proto-Seediq (See Section **??** for the alternations in Proto-Seediq).

This section attempts to explore all alternations as comprehensively as possible. However, there may still be areas that are not fully developed, such as specific alternations in subdialects of certain villages. In this section, segments participating in alternations and currently under discussion will be bolded, while alternations that are not the immediate focus will not be highlighted prominently.

In this section, if there is no difference between the phenomena in the Central

and Eastern varieties of Toda and Truku, only one example will be chosen from each, typically Central Toda and Eastern Truku, to represent them.

3.2.1 Vowel alternations

3.2.1.1 Prepenultimate vowel neutralizing

As discussed in Section **??**, all Seediq dialects exhibit the phenomenon of prepenultimate vowel neutralization. This phenomenon can also be referred to as weakening (Lee 2018: 22–23). The neutralizing vowel in Tgdaya is /u/, while in other dialects, it is /ə/.

This rule applies in morphological processes as well. For example, when adding monosyllabic suffixes such as *-un* 'PV' or *-an* 'LV', the original penultimate syllable becomes antepenultimate and thus neutralizes. Some suffixes consist of two syllables, such as *-ani* 'IMP.CV'. If the root originally has two syllables, then both vowels of the two syllables undergo neutralization. Refer to Table **??** for relevant examples in Central Toda, Eastern Truku, and Tgdaya.

Table 3.12: Vowel neutralization in Seediq dialects

Dialect	Bare form	Suffix	Suffixed form	Gloss
Central Toda	c**i**kul	-un	c**ə**kulun	'to push'
Eastern Truku	b**ey**taq	-an	b**ə**taqan	'to stab'
Tgdaya	p**ati**s	-ani	p**utu**sani	'to write'

In native words, a few exceptions have been discussed in the previous section. However, in modern loanwords borrowed from Japanese and Taiwanese, most do not

follow this rule. Table **??** presents those examples where the rule applies. Particularly noteworthy is the word for 'television', which in the Tgdaya dialect is *terebi*, where the rule does not apply, but in Central Toda *təreybi* and Central Truku *təribi*, it does.

Table 3.13: Vowel neutralization in loanwords

Dialect	Example	Source	Gloss
Tgdaya	c**u**bukiŋ	ts**a**-bóo-king (Taiwanese)	'brothel' > 'prostitute'
Central Toda	t**ə**reybi	t**e**rebi (Japanese)	'television'
Central Truku	t**ə**ribi	t**e**rebi (Japanese)	'television'

3.2.1.2 Alternations related to /ə/ or /e/

In Central Toda, Eastern Toda, Central Truku, and Eastern Truku, schwa [ə] is not allowed to occur in the final syllable, which is a quite stable phonetic restriction. In some words where the surface form ends in -*u*(C)# and a monosyllabic suffix is added, the [u] sound changes to [ə], which is related to the aforementioned restriction. Therefore, from a synchronic phonological perspective, [u] is an allophone of the phoneme /ə/ in the final syllable.

In the phonological system of Tgdaya, the sound [ə] is lacking, but we can observe a similar phenomenon, namely the alternation between [e] and [u]. The /e/ sound in Tgdaya originates partly from the historical diphthong *ay and partly from *ə. The alternation observed here is related to the historical schwa *ə. Section **??** will discuss the particular reflex of Proto-Seediq *ə in Tgdaya.

Many past studies have explored this phenomenon as well, such as Lee (2018), Sung (2018), and Yang (1976). Table **??** displays relevant examples from different di-

alects.

Table 3.14: Alternations related to /ə/ or /e/ in Tgdaya, Central Toda, and Eastern Truku

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	deŋ**u**	duŋ**e**un	'(to) dry'
Central Toda	kər**u**t	kər**ə**tun	'to cut'
Eastern Truku	rəq**u**n	rəq**ə**nun	'to swallow'

However, in the Toda and Truku dialects, there is a phenomenon involving opacity, namely the surface alternation between [u] and [a/i/(u)]. Such a phenomenon is also related to the (historical) schwa, but it actually involves vowel harmony, so it will be discussed in Section **??**.

3.2.1.3 Alternations related to vowel hiatus *-a.i-* and *-a.u-*

The structure such as C*a.i*C or C*a.u*C in the stem generates a special alternation when a monosyllabic suffix is attached. Taking the suffix *-un* 'PV' as an example, in Tgdaya, C*a.i*C-un and C*a.u*C-un respectively become C*e.Cun* and C*o.Cun*, in Central Toda, Central Truku, and Eastern Truku dialects they become C*ey.Cun* and C*ow.Cun*, **while in the Eastern Toda dialect they become C*ay.Cun* and C*aw.Cun* (ay 要再確認)**. This means that the two vowels originally belonging to different syllables merge when adding a monosyllabic suffix, leading to monophthongization or the formation of a diphthong in different dialects. Table **??** illustrates examples from Tgdaya, Eastern Truku, and Eastern Toda.

However, there is one exception: the word "to sew" in Tgdaya and Eastern Truku has a bare form of *sais*, but when the suffix *-un* 'PV' is attached, instead of the expected

Table 3.15: Alternations of related to vowel hiatus *-a.i-* and *-a.u-*

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	tu.nu.n**a.i**s	tu.nu.n**e**.sun	'rapidly'
Eastern Truku	pə.n**a.i**s	pə.n**ey**.sun	'rapidly'
Eastern Toda	問	問	
Tgdaya	t**a.u**s	t**o**.sun	'to wave hand'
Eastern Truku	t**a.u**s	t**ow**.sun	'to wave hand'
Eastern Toda	p**a.u**x	p**aw**.xi	'to turn over' (Lee 2015: 98)

sesun and **seysun, it becomes *si.i.sun* and *sə.i.sun*, respectively. This may be related to the historical development of Seediq and cannot be simply viewed as a synchronic phenomenon. The special case of this word will be revisited from a diachronic perspective in Section **??.

3.2.1.4 Vowel coalescence

The environment for vowel coalescence in Seediq dialects is quite consistent across all dialects, with the exception of Central Toda and Eastern Toda having an additional environment compared to other dialects. In Sections **??** and **??**, I have already discussed the occurrence of vowel sequences in the antepenultimate and penultimate syllables in both Toda dialects.

The common environment in other dialects is when a stem ends in a CV.V.CV(C) structure, and when a suffix is added, the original first two vowels, which belong to separate syllables, merge into one and are neutralized. Examples from each dialect and a comparison with Central Toda are shown in Table **??**.

Table 3.16: Vowel coalescence in Seediq dialects

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	pul**aa**le	pul**u**layun	
Eastern Truku	pəl**əa**lay	pəl**ə**layun	'to do first'
Central Toda	pəlalay	pələlayun	

3.2.1.5 Apharesis

The term "apharesis" (or "aphaeresis", < Greek *apo-* 'away' + *hairein* 'to take') refers to "A sound change in which a word-initial vowel is lost, as in the occasional pronunciation of *American* as '*merican*'." (Campbell and Mixco 2007: 13).

In Seediq, this phenomenon occurs extensively in all dialects. When there is a disyllabic stem and its penultimate syllable is onsetless, this phenomenon occurs. The aforementioned stem (VCV(C)) undergoes deletion when adding a suffix because the vowel has shifted to a non-stressed syllable position and lacks onset "protection". Table **??** shows relevant examples from Tgdaya, Central Toda, and Eastern Truku. In the table, the deleted vowels are represented by an underscore "_".

Table 3.17: Apharesis in Seediq dialects

Dialect	Bare form	STEM-un	STEM-ani	Gloss
Tgdaya	**e**yah	_yahun	_yahani	'to come'
Central Toda	**i**mah	_mahun	_məhani	'to drink'
Eastern Truku	**a**pa	_paun	_pəani	'to carry on back'

From the table above, it can be observed that regardless of whether monosyllabic

or disyllabic suffixes are added, when a vowel originally located in the penultimate syllable shifts to any prepenultimate syllable, it is deleted. However, there is one exception to this rule, which occurs during reduplication and may not apply uniformly across all dialects. For example, in the case of the Tgdaya word *alaŋ* 'village', its reduplicated form can be either *ula~alaŋ* or *la~alaŋ*, while in Eastern Truku it must be *ələ~alaŋ*.

3.2.1.6 Vowel harmony

The different dialects of Seediq all exhibit the phenomenon of vowel harmony, but Tgdaya and non-Tgdaya are almost entirely separable in terms of the patterns they involve.

In Tgdaya, vowel harmony typically occurs on the left side of the word. More precisely, this assimilation often takes place between the antepenultimate and penultimate syllables, as mentioned in Section **??**. When the penultimate syllable is onsetless or has an onset of *h*, the vowel of the antepenultimate syllable harmonizes with that of the penultimate syllable. This applies regardless of whether the antepenultimate vowel is originally in the word or when a bisyllabic stem is prefixed with a CV- prefix. See Yang (1976: 666–673) for more details about this issue. Table **??** illustrates the vowel harmonic patterns in Tgdaya. For ease of presentation, examples involving prefixation (*mu-* 'AV' and *su-* 'CV') are used in the table to show the differences.

In non-Tgdaya dialects, as mentioned earlier in Section **??**, there is an alternation involving opacity. This phenomenon is related to both historical schwa and vowel harmony. In these dialects, if the underlying /ə/ appears in the penultimate syllable and the final syllable is onsetless, the penultimate schwa is assimilated to final vowel

Table 3.18: Vowel harmonic patterns in Tgdaya

Dialect	Bare form	Prefixed form	Gloss
Tgdaya	apa	s**a**apa	‘to carry on back (CV)’
Tgdaya	imah	s**i**imah	‘to drink (CV)’
Tgdaya	hulis	s**u**hulis	‘to laugh (CV)’
Tgdaya	hediq	m**e**hediq	‘to give way (AV)’
Tgdaya	oda	m**o**oda	‘to go through (AV)’

following it. In other words, when suffixes are added to stems containing final /ə/ (all suffixes in Seediq begin with a vowel), the schwa is assimilated to the following vowel, leading to a certain degree of opacity. Table **??** compares Eastern Truku with Tgdaya.

Table 3.19: Alternations of [u] and other vowels in Eastern Truku and a comparison with Tgdaya

Dialect	Bare form	STEM-i	STEM-un	STEM-an	Gloss
Eastern Truku	dəŋ**u** /dəŋə/	dəŋ**i**i	dəŋ**u**un	dəŋ**a**an	‘to dry’
Tgdaya	deŋ**u** /deŋe/	duŋ**e**i	duŋ**e**un	duŋ**e**an	‘to dry’

3.2.1.7 Alternation between monophthongs and diphthongs (“VG” sequences)

This phenomenon is only observed in Tgdaya. Stems containing *-o* and *-e* (partially) exhibit alternations with *-a.w-* and *-a.y-* after adding suffixes. The alternation between *-e* and *-u-* has been discussed earlier in Section **??**, while the alternation involving *-o* and *-a.g-* will be introduced in Section **??**. Relevant examples can be found in Table **??**.

Table 3.20: Alternation between monophthongs and diphthongs in Tgdaya

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	su.lu.h**e**	su.lu.h**a.y**un	'to learn'
Tgdaya	ku.ka.r**o**	ku.ku.r**a.w**an	'to climb'

3.2.2 Consonant alternations

Consonantal alternations are highly correlated with the word-final position of the sound (with a few exceptions), often occurring when attaching suffixes due to no longer being in such position. However, sometimes a set of alternations may be related to more than one phonological rule, making classification challenging. Therefore, this section will provide separate explanations through direct examples and combine similar cases for discussion.

3.2.2.1 Alternations between *-c* (or *-t*) and *-d/t-*

This alternation occurs in Tgdaya, Central Toda, Eastern Toda, and Central Truku because /d/ and /t/ are not allowed in word-final positions and both are affricated (and devoiced) as *c* [ʦ]. Roots containing final /d/ or /t/ in non-suffixed forms meet these conditions, resulting in the alternation of word-final /d/ or /t/ to *c* [ʦ]. However, when suffixes are attached, [d] and [t] are allowed to appear since they are no longer in word-final positions, leading to this alternation. Please refer to Table **??** for relevant examples.

In Eastern Truku, word-final /d/ and /t/ merge into *-t* (though phonetically, both *-t* and *-c* may appear, but the orthography keeps the more conservative form). Thus, we only observe the alternation between *-d-* and *-t*. For example, *haŋu**t*** and *həŋə**d**un*

Table 3.21: *-d/t-* and *-c* alternations in Seediq dialects

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	haŋu**c**	huŋe**d**un	'to cook'
Central Toda	qiyu**c**	qiyu**t**un	'to bite'
Central Truku	lutu**c**	lətu**d**un	'to join two things'

Table 3.22: *-k* and *-b/p-* alternations in Seediq dialects

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	elu**k**	le**b**un	'to close'
Central Toda	adu**k**	du**p**an	'to hunt by chasing'
Eastern Truku	qari**k**	qəri**b**un	'to cut (w/ scissors)'

'to cook'.

3.2.2.2 Alternations between *-b/p-* and *-k*

This alternation also arises due to phonological constraints on word-final positions of the Seediq language, namely, (1) bilabial sounds are not allowed in word-final positions, and (2) voiced stops are not allowed in word-final positions. Consequently, the underlying forms of /b/ and /p/ become and merge as the velar [k]. Similarly, when suffixes are attached to roots ending in /b/ or /p/, the aforementioned environment disappears, and thus the original sound remains unchanged, resulting in alternation between different forms. This alternation occurs in all Seediq dialects. Table **??** shows the examples from different dialects.

3.2.2.3 Alternation between *-ŋ* and *-m-*

This alternation, similar to the previous one, arises from the phonological constraint that labial sounds cannot appear in word-final positions and are replaced by their velar counterparts. However, in this case, it involves nasal sounds. This alternation occurs in all Seediq dialects as well. See Table **??** for the examples.

Table 3.23: *-m-* and *-ŋ* alternations in Seediq dialects

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	geru**ŋ**	gure**m**un	'to break'
Central Toda	səməsu**ŋ**	səməsə**m**an	'to celebrate'
Eastern Truku	lau**ŋ**	low**m**un	'to burn'

3.2.2.4 Alternations related to *-Vg-* sequences

This type of alternation does not apply to the Central Toda and Eastern Toda dialects because they do not have the phoneme /g/ or /ɣ/ in their phoneme inventories. In the remaining dialects, the changes involving the sound *g* are quite complex, especially in the Tgdaya dialect. Yang (1976: 653–57) discusses the phenomenon using rule-based phonology, and Song and Pan (2024) analyze this set of alternations from the perspective of Optimality Theory. Here, I will only examine the surface alternation, as the goal is solely to understand its patterns. And because the Tgdaya dialect differs significantly from the two Truku dialects (with Eastern Truku as a representative here, since they behave identically), they will be discussed separately.

The sound /g/ is not permitted to appear at the end of words in the Tgdaya dialect.

However, unlike other voiced stops which directly undergo devoicing, /g/ tends to undergo an alternation with the preceding vowel. In Tgdaya, we observe the following four types of alternations: *ag ~ o*, *ig ~ uy*, *ug ~ u*, and *eg ~ u*. The vowel alternations in the last two cases are related to the phenomenon mentioned in Section **??**.

Table 3.24: Alternations related to -V*g*- in Tgdaya

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	reŋ**o**	ruŋ**ag**un	'to speak'
	bar**uy**	bur**ig**un	'to buy'
	tab**u**	tub**ug**un	'to raise (animal)'
	sep**u**	sup**eg**un	'to count'

In Eastern Truku, the situation is much simpler. In fact, we only observe the alternation between *aw* and *ag* [aɣ], as seen in examples like *rəŋ**aw** ~ rəŋ**ag**un* 'to speak'. However, in more recent variaties of the Truku dialects, *-ig* and *-ug* has changed to *-iy* and *-uw*, respectively, leading to the emergence of more alternation patterns. Table **??** illustrates these alternations in the newer variety of the Eastern Truku dialect.

Table 3.25: Alternations related to -V*g*- in Eastern Truku

Dialect	Bare form	Suffixed form	Gloss
Eastern Truku (Newer variety)	rəŋ**aw**	rəŋ**ag**un	'to speak'
	bar**iy**	bər**ig**un	'to buy'
	tab**uw**	təb**ug**un	'to raise (animal)'
	səp**uw**	səp**əg**un	'to count'

3.2.2.5 Alternations between *-n* and *-l/r-*

This alternation is exemplified by Tgdaya, as it represents the most complete set of changes associated with this phenomenon. As mentioned in Section **??**, in the Tgdaya dialect, most native speakers have merged the word-final *-l* and *-r* into *-n*. However, if stems containing final *-l/r* are attached with a suffix, the target segment is no longer in the word-final position, causing the original *l* or *r* to reappear. Examples related to this phenomenon are shown in Table **??**.

Table 3.26: *-n* and *-l/r-* ~ alternations in Tgdaya

Dialect	Bare form	Suffixed form	Gloss
Tgdaya	ciku**n**	cuku**l**un	'to push'
Tgdaya	aŋa**n**	ŋa**l**un	'to take'
Tgdaya	taku**n**	tunuku**r**an	'to fall'
Tgdaya	sudaha**n**	suduha**r**an	'to lean on'

3.2.2.6 Alternations between *p/b-* and *m-*

This alternation is one of the few regular dissimilation phenomena in Seediq. It occurs when a stem beginning with *p* or *b* takes the infix <um> (TG)/<əm> (other dialects) 'AV'. For example, in the Eastern Truku dialect, the verb root *patas* 'to write', when attached with the suffix <əm>, would be expected to form **pəmatas. However, due to both *p* and *m* being bilabial consonants, this leads to a dissimilation resulting in the deletion of the entire syllable "*pə-*". Then the surface form of *patas* + <*əm*> is thus *matas*, so, it appears as if the initial *p-/b-* is directly replaced by *m-* by looking

from the surface. Examples from Tgdaya, Central Toda, and Eastern Truku showing this phenomenon are presented in Table **??**.

Table 3.27: *p*/*b*- ~ *m*- alternations in Seediq dialects

Dialect	Bare form	Infixed form	Gloss
Tgdaya	**b**ekuy	**m**ekuy	'to tie'
Central Toda	**b**ari	**m**ari	'to buy'
Eastern Truku	**p**atas	**m**atas	'to write'

3.2.2.7 Palatalization and affrication of dental/alveolar stops before *i* and *y*

This type of alternation can be observed in Eastern Toda, Central Truku, and Eastern Truku due to the characteristic of palatalization of the sounds /t/ and /d/ before the high front vowel /i/ in these dialects, often accompanied by affrication. For convenience, this phenomenon is referred to as palatalization throughout this thesis. We can further classify these alternations into three types based on their manifestations: (1) palatalization disappears after adding the infix, (2) palatalization appears at the left side of the word after adding a suffix, and (3) palatalization appears at the right side of the word after adding a suffix *-i* 'IMP.PV'.

In the first type, the verb root needs to be in the form of /tiCV(C)/ or /diCV(C)/, which become *ci*CVC and *ɟi*CVC on the surface. When we add the suffix <əm>, the environment that caused palatalization originally disappears, so /t/ and /d/ remain unchanged. For example, in Eastern Truku, *cinun* 'to weave' + <*əm*> 'AV' becomes *təminun* instead of **cəminun. Table **??** shows relevant examples in Eastern Toda, Central Truku, and Eastern Truku dialects.

Table 3.28: Alternations related to palatalization of dental/alveolar stops in Seediq dialects: Pattern 1

Dialect	Bare form	Infixed form	Gloss
Eastern Toda	**ɟ**iɟil	**d**əmiɟil	'to carry'
Central Truku	**c**iyu	**t**əmiyu	'to point'
Eastern Truku	**c**inun	**t**əminun	'to weave'

The second type involves verb roots in the form of /tVjV(C)/ or /dVjV(C)/, where the first V cannot be /i/ (as it would palatalize in any case). When such stems take a suffix, the first V undergoes neutralization and is assimilated to a homorganic *i* due to the following *y* /j/, resulting in palatalization of the preceding /t/ or /d/. However, such examples are extremely rare, with only one instance found so far, namely *dayaw* 'to help'. Table **??** illustrates examples from each dialect.

Table 3.29: Alternations related to palatalization of alveolar stops in Seediq dialects: Pattern 2

Dialect	Bare form	Suffixed form	Gloss
Eastern Toda	**d**ayaw	**j**iyawun	'to help'
Central Truku	**d**ayaw	**j**iyagun	'to help'
Eastern Truku	**d**ayaw	**j**iyagun	'to help'

The third type is much more common compared to the second type. Whenever a root ends in /t/ or /d/ in their underlying representations, palatalization occurs when the suffix *-i* 'IMP.PV' is attached. In the bare form of such roots or in other forms without suffixation, /t/ and /d/ may change to *c* or *t* in the word-final position depending on the dialect (see Section **??**). Relevant examples are shown in Table **??**.

Table 3.30: Alternations related to palatalization of alveolar stops in Seediq dialects: Pattern 3

Dialect	Bare form	Suffixed form	Gloss
Eastern Toda	haŋu**c** /d/	həŋə**ɟ**i	'to cook'
Central Truku	qiyu**c** /t/	qiyu**c**i	'to bite'
Eastern Truku	həmu**t** /t/	həmə**c**i	'to do sth. casually'

3.3 Interim summary

In this chapter, I conducted a phonological sketch of five Seediq dialects, introducing their consonant and vowel inventories, syllable structures, and phonotactic constraints in Section **??**.

Section **??** was inspired by Li's (1977) investigation into synchronic alternations in various Formosan languages and Goderich's (2020) study of synchronic alternations in different Atayal dialects. In this section, I compared the morphophonological alternation patterns present in each Seediq dialect. Li (1977) proposed that Atayalic languages exhibit the richest alternations among Formosan languages, with Paran (= Tgdaya) Seediq showing the highest diversity, a conclusion consistent with the findings of this thesis.

Through the comparison of synchronic phonological differences, I believe that a significant portion can be attributed to historical changes. Discrepancies in phoneme inventory, for instance, likely stem from historical mergers or splits. The relationship between phonotactic constraints and synchronic alternations is closely intertwined, as phonotactic constraints, especially those at the word-final position, often give rise to alternations. Consistent alternation patterns across most dialects may reflect those in

the proto-language, while variations among dialects are indicative of surface historical sound changes.

Based on the above materials, I will reconstruct the phonological system of Proto-Seediq in Chapter **??**.

Chapter 4

Proto-Seediq phonology

In this chapter, I will first present the phoneme inventory of Proto-Seediq. Subsequently, I will list out each reconstructed phoneme along with their reflexes and lexical examples in modern dialects. I will also show and explain irregular and/or inconsistent sound correspondences and discuss what proto-phonemes they should be reconstructed as. Additionally, the phonotactics of Proto-Seediq and morphophonological alternations that can be reconstructed to the proto-language level will be addressed.

4.1 Proto-Seediq phoneme inventory

Table **??** shows the consonant inventory of Proto-Seediq. There are 18 consonants in Proto-Seediq based on my reconstruction, including:

(i) six stops: *p, *t, *k, *q, *b, *d

(ii) one affricate: *c

(iii) four fricatives: *s, *x, *h, *g

(iv) three nasals: *m, *n, *ŋ

(v) two liquids: *l, *r

(vi) two glides: *w, *y

Table 4.1: Proto-Seediq consonant inventory

	Bilabial	Dental/Alveolar	Palatal	Velar	Uvular	Pharyngeal
Stop	*p *b	*t *d		*k	*q	
Affricate		*c				
Fricative		*s		*x *g		*h
Nasal	*m	*n		*ŋ		
Lateral		*l				
Tap/Flap		*r				
Glide	*w		*y			

The vowel inventory of Proto-Seediq is straightforward, as shown in Table **??**. There are four monophthongs in Proto-Seediq, which are *a, *i, *u, and *ə. Note that mid vowels [e] and [o] cannot be found in Proto-Seediq. There are three "diphthongs" (VG sequences) in Proto-Seediq, including *ay, *aw, and *uy.

Table 4.2: Proto-Seediq vowel inventory

	Front	Central	Back	"diphthongs"	
High	*i		*u	*uy	
Mid		*ə			
Low		*a		*ay	*aw

4.2 Reconstructed phonemes and their reflexes

This section reconstructs the proto-phonemes in Proto-Seediq by comparing segments from words collected from modern dialects. Additionally, I demonstrates the reflexes of these reconstructed phonemes in modern dialects across different positions and environments (if applicable).

4.2.1 Proto-Seediq consonants

4.2.1.1 Proto-Seediq obstruents

Reflexes of Proto-Seediq *p can be found in word-initial and medial positions in all dialects, as given in the reflexes of *pahuŋ 'gallbladder' and *səpi 'dream' in Table **??**. However, no Seediq dialects have a word-final *-p*, because there was a merger of word-final *-k and *-p as *-k from Proto-Atayalic to (Proto-)Seediq (Li 1981). Li also points out that this merger is natural because [p] and [k] share the feature [+grave].

Table 4.3: Reflexes of Proto-Seediq *p

	*p-	*-p-
PSed	***p**ahuŋ	*sə**p**i
Tg	**p**ahuŋ	se**p**i
CTo	**p**ahuŋ	sə**p**i
ETo		
CTr	**p**ahuŋ	sə**p**i
ETr	**p**ahuŋ	sə**p**i
Gloss	'gallbladder'	'dream'

In suffixed forms of Proto-Seediq words that contain a final *-k from the historical **-p, we can find the final consonant alternating between *k (in non-suffixed forms, like AV form and bare form) and *p (in suffixed forms such as PV form and LV form). Such as *miyu**k** ‘to blow (AV)’ and *yu**p**un ‘to blow (PV)’. I will provide a more detailed description of the alternations in Proto-Seediq in Section **??**.

Reflexes of Proto-Seediq *t are relatively more complicated, as shown in Table **??**. Proto-Seediq *t is reflected as *c* before *i* in ETo, CTr, and ETr. More precisely, *t underwent a palatalization and an affrication processes and changed to an affricate *c* [tɕ] in this environment. The relevant examples can be found in the reflexes of *timu ‘salt’ and *quti ‘feces’ in above-mentioned dialects.

Table 4.4: Reflexes of Proto-Seediq *t

	*t-	*t- /_i	*-t-	*-t- /_i	*-t
PSed	***t**unux	***t**imu	*u**t**ux	*qu**t**i	*səpat
Tg	**t**unux	**t**imu	u**t**ux	qu**t**i	sepac
CTo	**t**unux	**t**imu	u**t**ux	qu**t**i	səpac
ETo					
CTr	**t**unux	**c**imu	u**t**ux	qu**c**i	səpac
ETr	**t**unux	**c**imu	u**t**ux	qu**c**i	səpat
Gloss	‘head’	‘salt’	‘spirit’	‘feces’	‘four’

In word-final position, Proto-Seediq *t is reflected as *c* [ʦ(ʰ)] in Tgdaya, Central Toda, Eastern Toda, and Central Truku. However, the final affrication seems to be optional and phonetic because the speakers do not release the final stops all the time,

especially if there is another word following it. For example, *sepac papak* 'four legs' (TG) can be pronounced as either [sepatsʰ papakʰ] or [sepatˀ papakʰ].

In addition to the environments mentioned above, Proto-Seediq *p is consistently reflected as /t/ in other positions, as seen in the reflexes of *tunux 'head' and *utux 'spirit' in Table **??**.

Proto-Seediq *k is reflected as /k/ in all positions in every dialect, as shown in the reflexes of *kari 'language', *yaku '1SG.NOM', and *papak 'animal foot' in Table **??**.

Table 4.5: Reflexes of Proto-Seediq *k

	*k-	*-k-	*-k
PSED	***k**ari	*ya**k**u	*papa**k**
TG	**k**ari	ya**k**u	papa**k**
CTO	**k**ari	ya**k**u	papa**k**
ETO			
CTR	**k**ari	ya**k**u	papa**k**
ETR	**k**ari	ya**k**u	papa**k**
Gloss	'language'	'1SG.NOM'	'animal foot'

The voiceless uvular stop *q is also regular in all the Seediq dialects under study, and can be found in all positions, as seen in reflexes of *quyux 'rain', *laqi 'child', and *dawriq 'eye' Table **??**. Li (1981) notes that Proto-Atayalic *q has shifted to an epiglottal (or pharyngealized) stop /ʔ/ in Truwan Seediq.

Reflexes of *b are generally regular in every dialect in the word-initial and medial positions, as shown in in Table **??**. Like /p/, /b/ can not be found in word-final position

Table 4.6: Reflexes of Proto-Seediq *q

	*q-	*-q-	*-q
PSed	***q**uyux	*la**q**i	*dawri**q**
Tg	**q**uyux	la**q**i	dori**q**
CTo	**q**uyux	la**q**i	dowri**q**
ETo			dawri**q**
CTr	**q**uyux	la**q**i	dowri**q**
ETr	**q**uyux	la**q**i	dowri**q**
Gloss	'rain'	'child'	'eye'

in the modern dialects either. *-b merged with *-k word-finally, but the final consonant alternation can still be noticed between the AV forms and LV/PV forms, as in *məlu**k** 'to close (AV)' and *lə**b**an 'to close (LV)'.

Table 4.7: Reflexes of Proto-Seediq *b

	*b-	*-b-
PSed	***b**irat	*u**b**al
Tg	**b**irac	u**b**an
CTo	**b**irac	u**b**al
ETo		
CTr	**b**irac	u**b**al
ETr	**b**irat	u**b**al
Gloss	'ear'	'fur; body hair'

Proto-Seediq *d is reflected as ʃ [ʒ~ʤ] in Eastern Toda, Central Truku, and Eastern

Truku before *i*, which parallels the development of Proto-Seediq *t. In other environments, Proto-Seediq *d is reflected as /d/ in all other word-initial and medial positions. Note that no word-final /d/ is found in any dialect. For reflexes of Proto-Seediq *d, please refer to *dara 'blood', *diyan 'daytime', *rudan 'old (person)', and *budi 'bow and arrow' in Table **??**.

Table 4.8: Reflexes of Proto-Seediq *d

	*d-	*d- /_i	*-d-	*-d- /_i
PSed	***d**ara	***d**iyan	*ru**d**an	*bu**d**i
Tg	**d**ara	**d**iyan	ru**d**an	bu**d**i
CTo	**d**ara	**d**iyan	ru**d**an	bu**d**i
ETo				
CTr	**d**ara	**ɟ**iyan	ru**d**an	bu**ɟ**i
ETr	**d**ara	**ɟ**iyan	ru**d**an	bu**ɟ**i
Gloss	'blood'	'daytime'	'old (person)'	'bow and arrow'

There is only one affricate *c in Proto-Seediq; its reflexes is shown in Table **??**, such as those of *cakus 'camphor' and *bicur 'worm'. Proto-Seediq *c is retained as an affricate *c* in Tgdaya and Central Toda, but it is reflected as *s* in Eastern Toda, Central Truku, and Eastern Truku. In addition, no word-final *c can be found in Proto-Seediq.

For Proto-Seediq *s, it is reflected as *s* in all positions in all five dialects under study, as shown in the reflexes of *sapah 'house', *misan 'winter', *cakus 'camphor' in Table **??**.

Proto-Seediq is regularly reflected as *x* in all positions and all dialects as well, as

Table 4.9: Reflexes of Proto-Seediq *c

	*c-	*-c-
PSed	***c**akus	*bi**c**ur
Tg	**c**akus	bi**c**un
CTo	**c**akus	bi**c**ur
ETo		
CTr	**s**akus	bi**s**ur
ETr	**s**akus	bi**s**ur
Gloss	'camphor'	'worm'

Table 4.10: Reflexes of Proto-Seediq *s

	*s-	*-s-	*-s
PSed	***s**apah	*mi**s**an	*caku**s**
Tg	**s**apah	mi**s**an	caku**s**
CTo	**s**apah	mi**s**an	caku**s**
ETo			
CTr	**s**apah	mi**s**an	saku**s**
ETr	**s**apah	mi**s**an	saku**s**
Gloss	'house'	'winter'	'camphor'

shown in the reflexes of *in Table **??**.

Table 4.11: Reflexes of Proto-Seediq *x

	*x-	*-x-	*-x
PSed	***x**iluy	*tə**x**al	*tunu**x**
Tg	**x**iluy	te**x**an	tunu**x**
CTo	**x**iluy	tə**x**al	tunu**x**
ETo			
CTr	**x**iluy	tə**x**al	tunu**x**
ETr	**x**iluy	tə**x**al	tunu**x**
Gloss	'iron'	'once'	'head'

However, the occurrence of *x is highly restricted. Word-initial *x appears only in one word — *xiluy 'iron'. This word appear to be a Seediq innovation, since there is no similar form found in other Formosan languages. Words that contain word-medial *x are also rare, they are usually related to the meaning 'one', including *təxal 'once', *taxa 'one person', *maxal 'ten', etc. These words all contain the fossilized sequences *-xa- or *-əxa-, which are supposed to be related to PAn *əsa 'one'. In Proto-Seediq, word-final *x can be found in quite a number of words.

Li (1981) mentioned that the historical source of this (proto-)phoneme is not yet clear, and Song (2024a) argues that (Proto-)Atayalic /x/ is borrowed from other neighboring languages, such as Pazeh and Saysiyat.

The regular reflex of Proto-Seediq *h is *h* in all positions in all dialects under study, as shown in the reflexes of *hidaw 'sun', *quhiŋ 'head louse', and *sapah 'house'

in Table ??.

Table 4.12: Reflexes of Proto-Seediq *h

	*h-	*-h-	*-h
PSed	***h**idaw	*qu**h**iŋ	*sapa**h**
Tg	**h**ido	qu**h**iŋ	sapa**h**
CTo	**h**idaw	qu**h**iŋ	sapa**h**
ETo			
CTr	**h**idaw	qu**h**iŋ	sapa**h**
ETr	**h**idaw	qu**h**iŋ	sapa**h**
Gloss	'sun'	'head louse'	'house'

Reflexes of Proto-Seediq *g are also complicated in the modern dialects, as shown in *gaya 'tradition; law'[1], *baga 'hand', *marig 'to buy (AV)', and *əlug 'road; path' in Table ??.

In word-initial and word-medial positions, Proto-Seediq *g is reflected as *w* in Central Toda and Eastern Toda, but it retains as *g* in all other dialects.

In the word-final position, the reflexes depend on the last vowel. Proto-Seediq *-g is retained as *g* after a high front vowel *i* in Central Truku and Eastern Truku; however, it is deleted in Central Toda and Eastern Toda (probably *-ig > **-iy > *i*). In Tgdaya, *-ig is reflected as *-uy*, which probably underwent the following gliding and metathesis processes: *-ig > **-iw > **-wi > *-uy*. As for *-g after *u, it is deleted (through a glide

[1]The Seediq term "*Gaya/Waya*" is difficult to find a direct equivalent in English or other languages like Mandarin and Japanese, so it is currently glossed as such as an expedient measure. "*Gaya/Waya*" may represent a collective set of norms for life, including what should or should not be done in certain situations. However, as I am not a member of the Seediq Nation, for a genuine interpretation of "*Gaya/Waya*", please refer to the understanding of Seediq people.

Table 4.13: Reflexes of Proto-Seediq *g

	*g-	*-g-	*-i**g**	*-u**g**
PSed	***g**aya*	*ba**g**a*	*mari**g**	*əlu**g**
Tg	**g**aya	ba**g**a	mar**u**y	elu
CTo	**w**aya	ba**w**a	mari	əlu
ETo				
CTr	**g**aya	ba**g**a	mari**g**	əlu**g**
ETr	**g**aya	ba**g**a	mari**g**	əlu**g**
Gloss	‘tradition; law’	‘hand’	‘to buy av)’	‘road; path’

/w/) in Tgdaya and Central Toda, but preserved as /g/ in Central Truku and Eastern Truku.

In the relatively early Tgdaya dialect, the word-final *-uw* was still present, as documented by Yang (1976) and Li (1981). However, the distinction between *-uw* and *-u* is no longer evident in the speech of my language consultant. The reason might be that the sequence [uw] violates the Obligatory Contour Principle (OCP), as [u] and [w] have nearly identical features, leading to further segment deletion.

Note that no evidence for the presence of word-final *-ag can be found in Seediq. It seems that Pre-Proto-Seediq **-ag merged with **-aw as *-aw in the Proto-Seediq stage. For example, Proto-Seediq *siy**aw** ‘next to’ and *bab**aw** ‘on, above’ correspond to Proto-Atayal *siy**ag** ‘rim; border; edge’ and *bab**aw** ‘on top of’, respectively.

There is a issue regarding the Proto-Seediq *g, whether this proto-phoneme should be classified as a stop or a fricative. In this paper, I choose to classify it as a fricative for the following reasons: (1) The reflexes of Proto-Seediq *g in modern dialects include [g], [w], [ɣ], indicating that the phonetic value of the proto-phoneme most likely was

[ɣ], (2) Seediq exhibits a phenomenon of word-final devoicing (from voiced stops to voiceless stops). However, whether in the Proto-Seediq stage or in modern dialects, there is no phenomenon where word-final *g or *g* changes to [k]. Instead, it interacts with different preceding vowels as a unit and becomes different sound sequences, without undergoing devoicing. Therefore, phonologically, *g is not classified in the same category as other voiced stop consonants such as *b and *d.

4.2.1.2 Proto-Seediq nasals

Proto-Seediq *m is reflected as *m* in word-initial and word-medial positions in all dialects, as seen in *malu 'good' and *tama 'father' in Table **??**.

Table 4.14: Reflexes of Proto-Seediq *m

	*m-	*-m-
PSed	***m**alu*	*ta**m**a*
TG	**m**alu	ta**m**a
CTo	**m**alu	ta**m**a
ETo		
CTr	**m**alu	ta**m**a
ETr	**m**alu	ta**m**a
Gloss	'good'	'father'

Word-final *-m* cannot be found in any dialects. Like other bilabial phonemes such as *p and *b, Proto-Seediq *-m merged with its corresponding velar phoneme *-ŋ. The alternation can still be found in AV forms vs. PV/LV forms, as contrasted in *migi**ŋ** 'to look for (AV)' and *gi**m**an 'to look for (LV)'.

Proto-Seediq *n and *ŋ are regularly reflected as *n* and *ŋ*, respectively in all positions in all dialects, as shown in the reflexes of *naqih 'bad', *inu 'where', and *rəbagan 'summer' in Table **??** and *ŋiraw 'mushroom', *məŋari 'nine', and *alaŋ 'village' in **??**.

Table 4.15: Reflexes of Proto-Seediq *n

	*n-	*-n-	*-n
PSed	***n**aqih	*i**n**u	rəbaga**n**
Tg	**n**aqah	i**n**u	rubaga**n**
CTo	**n**aqah	i**n**u	rəbawa**n**
ETo			
CTr	**n**aqih	i**n**u	rəbaga**n**
ETr	**n**aqih	i**n**u	rəbaga**n**
Gloss	'bad'	'where'	'summer

Table 4.16: Reflexes of Proto-Seediq *ŋ

	*ŋ-	*-ŋ-	*-ŋ
PSed	***ŋ**iraw	*məŋ**ŋ**ari	*ala**ŋ**
Tg	**ŋ**iro	mu**ŋ**ari	ala**ŋ**
CTo	**ŋ**iraw	məŋ**ŋ**ri	ala**ŋ**
ETo			
CTr	**ŋ**iraw	məŋ**ŋ**ari	ala**ŋ**
ETr	**ŋ**iraw	məŋ**ŋ**ari	ala**ŋ**
Gloss	'mushroom'	'nine'	'village'

4.2.1.3 Proto-Seediq liquids

Word-final *l is reflected as *n* in Tgdaya and is retained as *r* in other dialectl. The reflexes of *l in other position show regular *k*. Table **??** shows the reflexes of Proto-Seediq *l, as in the reflexes of *laqi 'child', *qawlit 'mouse', and *təxal 'once'.

Table 4.17: Reflexes of Proto-Seediq *l

	*l-	*-l-	*-l
PSed	***l**aqi	*qaw**l**it	*təxa**l**
Tg	**l**aqi	qo**l**ic	texa**n**
CTo	**l**aqi	qow**l**ic	təxa**l**
ETo			
CTr	**l**aqi	qow**l**ic	təxa**l**
ETr	**l**aqi	qow**l**it	təxa**l**
Gloss	'child'	'mouse'	'once'

The situation of Proto-Seediq *r is similar to that of *l. Tgdaya reflects word-final *r as *n*. Reflexes of *l in other positions in all dialects are *l*, as shown in the reflexes of *rudan 'old (person)', *dara 'blood', and *bicur 'worm' Table **??**.

4.2.1.4 Proto-Seediq glides

Proto-Seediq *w is reflected as *w* in word-initial and medial (syllable onset) positions in all dialects, as shown in the reflexes of *waqit 'fang', and *rawa 'basket' in Table **??**. Syllable-final *w will be discussed together with vowels in Section **??**.

Proto-Seediq word-initial and word-medial *y is reflected as *y* in all the dialects

Table 4.18: Reflexes of Proto-Seediq *r

	*r-	*-r-	*-r
PSed	***r**udan	*da**r**a	*bicu**r**
Tg	**r**udan	da**r**a	bicu**n**
CTo	**r**udan	da**r**a	bicu**r**
ETo			
CTr	**r**udan	da**r**a	bisu**r**
ETr	**r**udan	da**r**a	bisu**r**
Gloss	'old (person)'	'blood'	'worm'

Table 4.19: Reflexes of Proto-Seediq *w

	*w-	*-w-
PSed	***w**aqit	*ra**w**a
Tg	**w**aqic	ra**w**a
CTo	**w**aqic	ra**w**a
ETo		
CTr	**w**aqic	ra**w**a
ETr	**w**aqit	ra**w**a
Gloss	'fang'	'basket'

under study, as shown in the reflexes of *yaku ‘1SG.NOM’ and *gaya ‘tradition; law’. Syllable-final *y will be discussed together with vowels in Section **??**.

Table 4.20: Reflexes of Proto-Seediq *y

	*y-	*-y-
PSED	***y**aku	*ga**y**a
TG	**y**aku	ga**y**a
CTO	**y**aku	wa**y**a
ETO		
CTR	**y**aku	ga**y**a
ETR	**y**aku	ga**y**a
Gloss	‘1SG.NOM’	‘tradition; law’

4.2.2 Proto-Seediq vowels

4.2.2.1 Proto-Seediq monophthongs

Proto-Seediq *a, *i and *u are reflected as *a*, *i*, and *u*, respectively in all dialects, as shown in the reflexes of *alaŋ ‘village’, *ini ‘NEG’, and *rudu ‘nest’ in Table **??**. These three vowels can only be found in the penultimate syllable and the final syllable in Proto-Seediq.

As for Proto-Seediq *ə, it can only appears at a non-word-final position. However, the penultimate schwa and the prepenultimate ones behave very differently among modern dialects. Therefore, they will be discussed separately.

The penultimate *ə is regularly reflected as a mid front vowel *e* in Tgdaya, but

Table 4.21: Reflexes of Proto-Seediq *a, *i, and *u

	*a	*i	*u
PSed	***a**laŋ	***i**ni	***u**du
Tg	**a**laŋ	**i**ni	**u**du
CTo	**a**laŋ	**i**ni	**u**du
ETo			
CTr	**a**laŋ	**i**ni	**u**du
ETr	**a**laŋ	**i**ni	**u**du
Gloss	'village'	'neg'	'nest'

it is generally retained as *ə* in Central Toda, Eastern Toda, Central Truku, and Eastern Truku. In non-Tgdaya dialects, if the schwa *ə occurs before *y* or w, it will be assimilated as *i* or u. Moreover, if there is no consonant between the *ə and the next vowel, the *ə will be assimilated as the same value of the next vowel. Central Toda **(and Eastern Toda?)** tends to reflect *ə as *u* when there is a *g (> *w* in Central Toda) before or after it, but this does not occur in every case. In addition, if the historical *g occurs in the onset position, it will be deleted after the assimilation process occurs. Table **??** shows the reflexes of Proto-Seediq penultimate *ə as the reflexes of *beyax 'power', *wəəwa 'young woman', *məhəal 'to carry on shoulder', *ləiŋ 'to hide', *gesak 'k.o. bamboo tool', *məgay 'to give', and *əlug 'road; path' in modern dialects. The environments listed in the table are all situations where schwa *ə* appears in a penultimate syllable.

Unlike some Atayal dialects (e.g. Matu'uwal and Plngawan, see Goderich 2020 for details), none of the modern Seediq dialects still have retained phonemic distinctions of vowel in prepenultimate syllables. Generally speaking, it is better to reconstruct *ə

Table 4.22: Reflexes of Proto-Seediq penultimate *ə

	*ə /_y	*ə /_w	*ə /_V_x	
PSed	*b**ə**yax	*w**əə**wa	*m**əhə**al	*l**əi**ŋ
Tg	b**e**yax	we**e**wa	meh**e**an	l**e**iŋ
CTo	b**i**yax	w**əu**wa/wowwa	məhal	(tə)liŋ
ETo				
CTr	b**i**yax	wa**u**wa	məh**aa**l	l**ii**ŋ
ETr	b**i**yax	wa**u**wa	məh**aa**l	l**ii**ŋ
Gloss	‘power’	‘young woman’	‘to carry on shoulder’	‘to hide’

	*ə /g_	*ə /_g	*ə
PSed	*g**ə**sak	*m**ə**gay	***ə**lug
Tg	g**e**sak	m**e**ge	**e**lu
CTo	**u**sak	m**u**way	**ə**lu
ETo			
CTr	g**i**sak	m**ə**gay	**ə**lug
ETr	g**ə**sak	m**ə**gay	**ə**lug
Gloss	‘k.o. bamboo tool’	‘to give’	‘road; path’

for this prepenultimate neutralized (or reduced) vowel. It is reflected as *ə* in Central Toda, Eastern Toda, Central Truku, and Eastern Truku, and *u* in Tgdaya if there is no other condition, as seen in the reflexes of *dəqəras 'face' in Table **??**. [u] is more marked compare with [ə], and only one dialect — Tgdaya reflects *u* because [ə] is totally absent in its phonology; thus, *ə is assigned because of the above reasons.

As for the conditioned reflexes, again, in non-Tgdaya dialects, if the schwa *ə occurs before *y*, it will be assimilated as *i*. Moreover, since Central Toda and Eastern Toda reflect *w* of Proto-Seediq *g, it affected the vowel quality of the prepenultimate *ə. Specifically, the prepenultimate *ə is reflected as *u* in Central Toda if there is a *g before or after it. In addition, if the historical *g occurs in the onset position, it will be deleted after the assimilation process occurs. Relevant cases can be seen in the reflexes of *təyaquŋ 'crow', *gəciluŋ 'sea' and *bəgihur 'wind' in Table **??**.

In the Tgdaya dialect, the prepenultimate vowel will be assimilated by the following vowel if: (1) there is no consonant between two vowels, (2) the consonant between them is an *h* [ħ]. The reflexes of *məuyas 'to sing' and *cəhədil 'heavy' in Table **??** shows the examples.

Note that the environments listed in Table **??** are all situations where schwa *ə* appears in a **pre**penultimate syllables.

In addition, Central Toda, Eastern Toda and some Central Truku speakers tend to "shorten" a sequence of two identical vowels, as shown in the reflexes of *səədiq 'human' and *tələətu 'cold (thing)' in Table **??**. Note that this is mandatory in Central Toda **and Eastern Toda**, but it seems not very stable in Central Truku.

Table 4.23: Reflexes of Proto-Seediq prepenultimate *ə

	*ə /_y	*ə /g_	*ə /_g	*ə /_(h)V_x		*ə
PSed	*t**ə**yaquŋ	*g**ə**ciluŋ	*b**ə**gihur	*m**ə**uyas	*c**ə**hədil	*d**ə**qəras
Tg	t**u**yaquŋ	r**u**ciluŋ	b**u**gihun	m**u**uyas	c**e**hedin	d**u**qeras
CTo	t**i**yaquŋ	**u**ciluŋ	b**u**wihur	muyas	c**ə**hədil	d**ə**qəras
ETo		**u**siluŋ				
CTr	c**i**yaquŋ	g**ə**siluŋ	b**ə**gihur	m**ə**uyas	s**ə**hə**ɟ**il	d**ə**qəras
ETr	c**i**yaquŋ	g**ə**siluŋ	b**ə**gihur	m**ə**uyas	s**ə**hə**ɟ**il	d**ə**qəras
Gloss	‘crow’	‘sea’	‘wind’	‘to sing’	‘heavy’	‘face’

Table 4.24: Reflexes of Proto-Seediq *V_xV_x sequence

	*V_xV_x	
PSed	*s**əə**diq	*t**ələ**tu
Tg	s**ee**diq	tul**ee**tu
CTo	s**ə**diq	t**ə**l**ə**tu
ETo	s**ə**ɟiq	
CTr	s**əə**ɟiq	t**ə**l**ə**tu
ETr	s**əə**ɟiq	(m**ə**)t**ələ**tu
Gloss	‘human’	‘cold (thing)’

4.2.2.2 Proto-Seediq diphthongs

Proto-Seediq *ay and *aw are reflected as *e* and *o*, respectively in word-medial and final positions in Tgdaya. By contrast, they are retained *-ay* and *-aw* word-finally, but as *-ey-* and *-ow-* word-medially in Central Toda, Central Truku, and Eastern Truku. Among the dialects, Eastern Toda in the most conservative in terms of diphthongs. Eastern Toda reflects *ay and *aw as *ay* and *aw* in both word-medial and word-final positions.

Proto-Seediq *uy can only appear at the word-final position, and it is reflected as *-uy* in all dialects. Reflexes of Proto-Seediq diphthongs are shown in the reflexes of *bayluh 'k.o. bean', *balay 'true', *dawriq 'eye', *babaw 'above; top', and *babuy 'pig' in Table ??.

Table 4.25: Reflexes of Proto-Seediq diphthongs

	*-ay-	*-ay	*-aw-	*-aw	*-uy
PSed	*b**ay**luh	*bal**ay**	*d**aw**riq	*bab**aw**	*bab**uy**
Tg	b**e**luh	bal**e**	d**o**riq	bob**o**	bab**uy**
CTo	b**ey**luh	bal**ay**	d**ow**riq	bab**aw**	bab**uy**
ETo			d**aw**riq		
CTr	b**ey**luh	bal**ay**	d**ow**riq	bab**aw**	bab**uy**
ETr	b**ey**luh	bal**ay**	d**ow**riq	bab**aw**	bab**uy**
Gloss	'k.o. bean'	'true'	'eye'	'above; top'	'pig'

Except for Eastern Toda, we can observe that some records from the Japanese colonial period indicate that various Seediq dialects retained the word-medial *-ay-* and

-aw-. For example, the Musha/Paran dialect (related and similar to modern Tgdaya) in Ogawa (2006) includes words like *d**ao**ryak* (22), *d**ao**deak* (23a) 'eye', *b**ai**roaχ* (23d) 'bean', etc.[2]

Similarly, Tashiro's (1900) investigation of the Eastern Truku dialect (Taroko, 太魯閣) contains words such as *t**au**kan* 'net bag', *b**ai**lo* 'k.o. bean' that retained *-ay-* and *-aw-* in their lexicon.

However, there are also instances where speakers or villages documented in these sources have shifted to *-e(y)-* and *-o(w)-*. The raising of these two diphthongs or their further change to monophthongs appears to be sound changes that occurred independently in different dialects over the past century. Therefore, the diphthongs in the word-medial position should be reconstructed as *-ay- and *-aw- in the Proto-Seediq stage and/or at any possible subgroup's ancestral language level (see Chapter **??** for subgrouping and the Appendix for the word list).

4.2.3 Inconsistent sound correspondences

4.2.3.1 Inconsistent consonantal sound correspondences

Seediq dialects exhibit some unstable or inconsistent consonant correspondences. These inconsistent consonant correspondences may have various causes. This section will explore some inconsistent consonant correspondences that can be grouped into categories.

The first category considered is the sometimes irregular *l* : *d* correspondences between Seediq dialects. Generally, we would expect to see either all *d* or all *l*. However, in the examples below in Table **??**, there are inconsistent reflexes among the dialects.

[2]The code in the parentheses represents forms collected from different speakers, villages, and/or dialects in the book.

It is hard to decide which protophoneme I should assign to these correspondences. I will temporarily place [?] as an placeholder for this unknown value.

Table 4.26: Examples of *l* : *d*

PSed	*ba[?]aw	*ə[?]uk	*[?]ayat	*kənəda[?]ax
Tg	bado	əluk	dayac	kunudalax
CTo	balaw	əluk	dayac	kənədadax
CTr	balaw	əduk	layac	kənədadax
ETr	balaw	əduk	layat	kənədadax
Gloss	'a thorny vine'	'to close'	'Sambucus formosana'	'from; since'

Reconstructing this set of correspondences to the proto-level is actually quite challenging. The reflexes as *d* or *l* are distributed across different dialects within each lexical item, as shown in Table **??** for 'to close' and 'Sambucus formosana', where the distributions are completely opposite. Therefore, this phenomenon is unlikely to originate from the same protophoneme, suggesting other possibilities.

In fact, a similar "split" can be seen in the reflexes of Proto-Austronesian *N in Proto-Seediq. For example, PAn *paNid > PSed *palit 'wing' and PAn *siNaR > PSed *hidaw 'sun' show this difference.

Dyen and Tsuchida (1987) were likely the first scholars to notice this phenomenon. They suggested adding a new PAn symbol, *D_5, to distinguish between two sets of reflexes in Seediq: PAn *N > Seediq *l* and PAn *D_5 > Seediq *d*[3]. Among all Formosan languages, only Seediq has maintained this distinction, while in others, it has merged with *N.

[3]See Tsuchida (1976) for the detailed discussion of PAn *$D_{1\text{-}4}$.

Song (2024c) reevaluated the two reflexes of Proto-Austronesian *q in Proto-Seediq, which are *q and ∅. This study concluded that ∅ is actually the regular reflex in Seediq, while *q /q/ was borrowed from Atayal or its early stages. Supporting evidence suggests that the instances where Proto-Austronesian *N is reflected as *d in Proto-Seediq generally align with the pattern of PAn *q > PSed ∅, as seen in PAn *qaNup > PSed *aduk 'to hunt (w/ dogs)', *qəNeb > *ə[d/l]uk? 'to close (door)', *qiNaS > *idas 'moon', etc. This viewpoint from Song (2024c) indirectly suggests the possibility that Seediq /l/ might also have spread from Atayal. However, the distribution of *l* and *d* does not appear to be as clear as that of *q* and ∅.

I believe that the additional PAn phoneme proposed by Dyen and Tsuchida (1987) is unnecessary because the phenomenon might be related to the stratification within Seediq based on discussion in Song (2024c). However, one point mentioned by Dyen and Tsuchida (1987) is worth considering: when there is an *l* : *d* correspondence between Seediq dialects, the *l* group likely comes from Atayal. Therefore, when encountering such cases, I have a reason to reconstruct these words as Proto-Seediq *d, with the reflex as *l* possibly being a loanword from Atayal.

Some examples do not have a clear borrowing source and are often unique to Seediq, as seen in Table **??** with *kunudalax* : *kənədadax* 'from; since', or the *ladiŋ* : *laliŋ* 'fishhook' mentioned by Dyen and Tsuchida (1987: 175). I speculate that /l/ might not be a native phoneme of Seediq, and that Seediq speakers learning this sound might produce phonetic errors, confusing it with the native /d/. However, this argument needs to be discussed and validated through further research in the future, and it remains a hypothesis for now. Alternatively, this might simply be cases of sporadic assimilation. For these examples, I will temporarily use *[d/l] as a possible reconstruction approach.

The revised reconstructions are shown in Table **??**.

Table 4.27: Examples of *l* : *d* with the reconstructions

PSed	*badaw	*əduk	*dayat	*kənəda[d/l]ax
Tg	bado	əluk	dayac	kunudalax
CTo	balaw	əluk	dayac	kənədadax
CTr	balaw	əduk	layac	kənədadax
ETr	balaw	əduk	layat	kənədadax
Gloss	'a thorny vine'	'to close'	'Sambucus formosana'	'from; since'

The second category is *t* : *l*, which occurs possibly when there is a *b* sound within the word. In the available forms, Tgdaya and Central Toda both show *t*, Eastern Toda consistently shows *l*, and Central Toda has instances of both. Based on external evidence from Proto-Austronesian and Atayal, it is indicated that these *t* sounds likely originated from a historical *l. However, it is difficult to rule out that in Seediq, *t* is the regular reflex in this specific environment, and the dialects reflecting *l* may have borrowed from Atayal. Therefore, I tentatively list both t and l in Proto-Seediq as *[t/l].

Cases in Proto-Seediq with such special correspondences are: *bə[t/l]uku 'flat basket; winnowing basket', *[t/l]ibu 'pigsty', *[t/l]ubug 'mouth harp; musical instruments', as shown in Table **??**.

The third group of examples found within Seediq, related to the *c*/*s* : *l* correspondence, includes only two cases: (1) *siŋas* : *liŋas* 'food between teeth' and (2) biciq : bilaq 'small'[4]. Based on the irregular reflections of Proto-Austronesian *C and *S in the Atayalic languages, I reconstruct these inconsistent correspondences to Proto-Seediq

[4]The sequence *-aq* might come from **-iq sporadically. See Section **??** for details.

Table 4.28: Examples of $t : l$

PAN	—	*Nibu	—
PATA	*baluku	*libu	*lubug
Proto-Seediq	*bə[t/l]uku	*[t/l]ibu	*[t/l]ubug
Tgdaya	butuku	tibu	tubu
Central Toda	bətuku	tibu	tubu
Central Truku	bəluku	libu	tubug
Eastern Truku		libu	lubug
Gloss	‘flat basket; winnowing basket’	‘pigsty’	‘mouth harp; musical instruments’

as *c and *s, rather than *l. Relevant examples are shown in Table **??**.

Table 4.29: Examples of $c/s : l$

PSED	*siŋas	*biciq
TG	siŋas	biciq
CTO	liŋas	
CTR	siŋas	
ETR	liŋas	bilaq
Gloss	‘food between teeth’	‘small’

I compared the reflexes of Proto-Austronesian *C and *S in Seediq and Atayal, and found that there are sometimes unexpected reflexes as *l. Such reflexes occur occasionally in Seediq and sometimes in Atayal. For example, the reflexes of Proto-Austronesian *Cabu ‘to wrap’, *CuSuR ‘to string’, *dakiS ‘to climb’, *RamiS ‘root’, and *SipuR ‘to count’ can be seen in Table **??**.

Table 4.30: Examples of irregular PAn *C/*S > Atayalic /l/

PAn	*Cabu	*CuSuR	*dakiS	*RamiS	*SipuR
PSed	***l**abu	***l**ihug	*daki**l**	*gami**l**	*səpug
PAta	*cabu	***l**Vhug	*rakiyas	*gami**l**	***l**əpug
Gloss	'to wrap'	'to string'	'to climb'	'root'	'to count'

This phenomenon can also be observed in Atayal (Goderich 2020: 173). Based on external evidence from Proto-Austronesian, I choose to reconstruct these unique correspondences as Proto-Seediq *c or *s. The change to /l/ is likely a later development, occurring multiple times (PAn/PAic > PSed and PSed > Seediq dialects). With the current evidence, it is difficult to ascertain why Atayalic languages exhibit such consonant substitutions. However, the direction of *c/*s > *l* is quite clear.

The last group of inconsistent correspondences is *g* (or Central Toda, Eastern Toda *w*) : *r*. I name this as the "GR problem" of Seediq. Sometimes, the sounds *g* and *r* randomly distributed in specific words and the process seems to be dialect-independent, as given in Table **??**. However, there is no environment of these correspondences that can be generalized.

Li (1981) mentioned that Proto-Atayalic *g is reflected as *r* in Seediq when adjacent to the vowel *a*. However, Song (2024b) pointed out that such irregular sound changes had already occurred once in the evolution from Proto-Austronesian or Proto-Atayalic to Proto-Seediq at the word-initial, word-medial and word-final positions, not limited to the environments found by Li (1981). The aforementioned unconditioned split remains difficult to describe systematically until now. Nevertheless, Song (2024b) used the reflexes of Proto-Atayalic or Pre-Proto-Seediq antepenultimate syllable (and also in the word-initial position) (*)*w- in Proto-Seediq and Seediq dialects as evidence,

Table 4.31: Examples of *g* (or *w*) : *r*

PSed	*gəhak	*təmaga	*hagat	*məgipuh
TG	rehak	tumara	harac	mugipuh
CTO	rəhak	təmawa	hawac	məripuh
CTR	gəhak	təmaga	harac	məripuh
ETR	gəhak	təmaga	hagat	məripuh
Gloss	‘seed’	‘to wait’	‘stone fence’	‘fragile; soft’

PSed	*gəciluŋ	*gupun	*gəbəyuk	*gəmimax
TG	ruciluŋ	rupun	rubeyuk	gumimax
CTO	wuciluŋ	rupun	rəbiyuk	mərimax
CTR	rəsiluŋ	gupun	—	—
ETR	gəsiluŋ	gupun	gəbiyuk	gəmimax
Gloss	‘sea; lake’	‘teeth’	‘canyon’	‘to mix’

explaining that *g* > *r*, though rare in the sound changes of world languages, is a more likely direction to appear in Seediq. Relevant examples are shown in Table **??**.

Table 4.32: PAta antepenultimate *w- : Seediq *g*/*r* (from Song 2024b: 13)

PAic	PSed	Tg	To	CTr	ETr	Gloss
*__w__aciluŋ	*__[g/r]__əciluŋ	__r__uciluŋ	uciluŋ	__g__əsiluŋ	__g__əsiluŋ	'sea'
*__w__aqanux	*__r__əqənux	__r__uqenux	__r__əqənux	__r__əqənux	__r__əqənux	'deer'

Song (2024b: 13) argues that the probability of **w directly becoming *r* is relatively low, and there must have been an intermediate stage involving **g. So, we have sufficiently compelling evidence to suggest that *r* may indeed from *g*. Moreover, some instances of Proto-Atayal *g > Atayal dialects *r* or *ɹ* may also suggest a similar sporadic sound change pattern in the Atayalic family.

Therefore, based on the findings of Song (2024b), I reconstruct the inconsistent *g* : *r* correspondences in Seediq dialects as Proto-Seediq *g. This approach is a compromise, but it is the more probable choice given the current evidence.

4.2.3.2 Inconsistent vocalic sound correspondences

Among modern dialects of Seediq, there are sometimes irregular sound correspondences. Although these correspondences are unusual, some patterns can be identified. Therefore, these irregular correspondences may arise from analogy or sporadic changes in certain environments.

The first group of irregular correspondences involves the Tgdaya *i* corresponding to *a* in other dialects. This includes the following examples: (1) *matis* : *matas* 'to write', (2) *madis* : *madas* 'to bring', (3) *gumabin* : *wumabal*/*gəmabal* 'to pull', as shown

in Table ??.

Table 4.33: Examples of sporadic Tgdaya i : other a

Env.	/[cor]_[cor]	/[cor]_[cor]	/_[cor]
Proto-Austronesian	*pataS ‘tattoo’	*adaS	
Proto-Seediq	*matas	*madas	*gəmabal
Tgdaya	matis	madis	gumabin
Central Toda	matas	madas	wumabal
Eastern Truku	matas	madas	gəmabal
Gloss	‘to write’	‘to bring’	‘to pull’

The environment for such a sound change is clear. The first two examples indicate that *a may more easily change to *i* in Tgdaya between two coronal consonants. Similarly, a similar change occurs when only the coda is a coronal consonant, as in *gəmabal > *gumabin*. Interestingly, a sporadic sound change has occurred once before with Proto-Austronesian *dapaN > Proto-Seediq *dapil ‘sole (foot), footprint’ between a bilabial stop and a coronal coda.

This sound change likely results from the influence of coronal features on the vowel, producing sporadic sound changes. Because there is external evidence, we can confidently reconstruct these correspondences to Proto-Seediq *a.

In the other set of cases, the correspondence is reversed, with Tgdaya (and sometimes Central Toda 問 **Eastern Toda**) *a* : others *i*. This correspondence occurs when the target vowel is surrounded by two uvular or pharyngeal consonants. Additionally, this may also occur as an *i* ~ *a* variation within the same dialect before fully changing

to the sound *a*. For example: (1) *naqah* : *naqih* ‘bad’, (2) *luqah* : *luqih* ‘wound’, (3) *nuqah* : *nuqih* ‘hemp fiber’, (4) *rumehaq* : *rəməhiq* ‘to peel bark’, etc., as detailed in Table **??**.

Table 4.34: Examples of sporadic *a* : *i*

Env.	/q_h	/q_h	/q_h	/h_q
Proto-Seediq	*naqih	*luqih	*nuqih	*rəməhiq
Tgdaya	naqah	luqah	nuqah	rumehaq
Central Toda	naqah	luqah	nuqih ~ nuqah	rəməhiq
Eastern Truku	naqih	luqih	nuqih	rəməhiq
Gloss	‘bad’	‘wound’	‘hemp fiber’	‘to peel bark’

These sporadic sound changes are likely due to the vowel lowering effect caused by uvular and pharyngeal consonants. Generally, the vowel lowering forms an allophone of /i/, and is accompanied by a transitional vowel, such as /seediq/ > [seedi^{e}q] [seedi^{a}q] ‘person’. The direct shift to a single [a] is relatively rare and is likely caused when both preceding and following consonants are either uvular or pharyngeal.

4.3 Proto-Seediq phonotactics

Table **??** presents all syllable types in Proto-Seediq, including V, CV, VC, VG, CVC, and CVG. Among them, VG and CVG are relatively special in nature, and these two categories are separately listed from VC and CVC.

The only syllable type that can occur in prepenultimate syllable(s) is CV, more precisely, Cə, as in *qə from *qə.si.ya ‘water’ in Table **??**. A standalone V in that position

Table 4.35: Proto-Seediq syllable types

Syl. type	Prepenultimate	Penultimate		Final	
V	—	*__i__.ma	‘who’	*sə.kə.lu.__i__	‘startled’
VC	—	—		*sa.__un__	‘to go (PV)’
CV	*__qə__.si.ya ‘water’	*qə.__si__.ya	‘water’	*qə.si.__ya__	‘water’
CG	—	*__aw__.raw	‘k.o. bamboo’	*də.mu.__uy__	‘to hold (AV)’
CVC	—	—		*ba.__raq__	‘lungs’
CVG	—	*__daw__.riq	‘eye’	*ru.__ŋay__	‘monkey’

is either deleted or undergoes vowel coalescence (which is to some extent a deletion), hence it does not appear.

In the penultimate syllable, V, CV, VG, and CVG are all allowed to occur, as seen in the examples from the table: *i.ma ‘who’ where *i appears, *si from *qə.si.ya ‘water’, *aw in *aw.raw ‘k.o. bamboo’, and *daw in *daw.riq ‘eye’. Of these, the *aw in *aw.raw ‘k.o. bamboo’ is particularly noteworthy as it is the only case of VG found in the penultimate position.

As for the restrictions in the final syllable, there are the least limitations, with examples for all six syllable types. For instance, *i in *sə.kə.lu.i ‘startled’, *un in *sa.un ‘to go (PV)’, *ya in *qə.si.ya ‘water’, *uy in *də.mu.uy ‘to hold (AV)’, *raq in *ba.raq ‘lungs’, and *ŋay in *ru.ŋay ‘monkey’.

In modern Seediq dialects, there are several segments that are not allowed to occur at the word-final position across all dialects. Since the performance of all dialects is consistent, I reconstruct these restrictions back to the level of Proto-Seediq. The sounds not permitted in the word-final position in Proto-Seediq include the voiced stops *b, *d, the affricate *c, the labials *p, (*b,) *m, and a specific VC sequence *-ag. The historical

schwa *ə also cannot occur in the final syllable.

These phonological constraints further lead to the occurrence of many morphophonological alternations. I will provide detailed explanations about these alternations in Section **??**. It is worth noting that the absence of *-c did not cause any alternation, unlike in Proto-Atayal (see Goderich 2020: 60–61, 138–139).

4.4 Proto-Seediq morphophonological alternations

As mentioned in Section **??**, many alternations observed in modern Seediq dialects can also be found in Proto-Seediq. Additionally, with the reconstructions in Section **??** and the discussion of Proto-Seediq phonotactics in Section **??**, it can be inferred that the alternations in Proto-Seediq are caused by the constraints imposed by phonotactics. In this subsection, I will introduce the morphophonological alternations in Proto-Seediq one by one.

4.4.1 Vowel alternations in Proto-Seediq

4.4.1.1 Prepenultimate vowel neutralizing (weakening) in Proto-Seediq

Since all modern Seediq dialects exhibit prepenultimate vowel neutralization, I also reconstruct it back to the level of Proto-Seediq, with its phonetic value being *ə (please refer to Section **??** for the reconstruction of prepenultimate *ə).

Therefore, if originally in the bare form, vowels with phonemic distinctiveness (those located in penultimate and final syllables) may all shift to the prepenultimate position when the base is attached with suffixes of different lengths such as *-un ‘PV’, *-an ‘LV’, or *-ani ‘CV.IMP’. This positional change results in the neutralization of these

vowels and subsequently gives rise to alternations between suffixed and non-suffixed forms. Relative examples are shown in *c**i**kul ~ *c**ə**kulun ‘to push’, *b**ay**taq ~ *b**ə**taqan ‘to stab’, and *p**a**t**a**s ~ *p**ə**t**ə**sani ‘to write’ in Table **??**.

Table 4.36: *V~*ə alternation in Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
*V(G)~*ə	*c**i**kul	*c**ə**kulun	‘to push’
	*b**ay**taq	*b**ə**taqan	‘to stab’
	*p**a**t**a**s	*p**ə**t**ə**sani	‘to write’

4.4.1.2 *u~*ə alternation in Proto-Seediq

The restriction against the occurrence of schwa /ə/ in the final syllable is a shared feature among all modern dialects of Seediq. Such a restriction can be reconstructed at the Proto-Seediq level. This phenomenon is also observed in Atayal, representing a common innovation in the Atayalic languages (Goderich 2020; Li 1981).

This restriction results in the underlying final-syllabic schwa *ə surface as *u. However, when adding monosyllabic suffixes like *-un ‘PV’ or *-an ‘LV’[5], the aforementioned condition no longer applies, meaning that the underlying *ə appears in the penultimate syllable and is realized as *ə on the surface. Therefore, the differential realization of *ə in the final and penultimate syllables gives rise to the alternation between *u and *ə, as shown in *həmaŋ**u**t ~ *həŋ**ə**dun ‘to cook’ and *kəmər**u**t ~ *kər**ə**tun ‘to saw’ in Table **??**.

[5]The voice affixes in Proto-Seediq will be elaborated in Section **??**.

Table 4.37: *u~*ə alternation in Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
*u~*ə	*həmaŋ**u**t	*həŋ**ə**dun	'to cook'
	*kəmər**u**t	*kər**ə**tun	'to saw'

4.4.1.3 Alternations related to vowel hiatus *-a.i- and *-a.u- in Proto-Seediq

In Section **??**, we discussed the resolution of hiatus in bare form structures like C*a.i*C or C*a.u*C when attached by monosyllabic suffixes, resulting in various patterns across different dialects such as C*e*.C*un*/C*o*.C*un*, C*ey*.C*un*/Cow.C*un*, and C*ay*.C*un*/C*aw*.C*un*. In fact, following the reconstruction of diphthongs in Section **??**, it can be observed that the correspondences of *e/o* : *ey/ow* : *ay/aw* in modern dialects are inherited from Proto-Seediq *-ay- and *-aw-. Therefore, these alternations in modern dialects are indeed remnants from Proto-Seediq. The examples of alternations between *-a.i-/*-a.u- and *-ay-/*-au- are shown in *n**ai**s ~ *n**ay**sun 'rapidly' and *t**au**s ~ *t**aw**sun 'to wave hand' in Table **??**.

Table 4.38: *u~*ə alternation in Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
*a.i ~*ay	*n**ai**s	*n**ay**sun	'rapidly'
*a.u ~*aw	*t**au**s	*t**aw**sun	'to wave hand'

We also discussed a special case in Section **??**, the word 'to sew', which can be reconstructed in Proto-Seediq as the bare form *sais and the suffixed form *səisun. However, following the alternation pattern mentioned above, the expected suffixed

form should be **saysun.

One possible explanation is that *sais in Proto-Seediq is a loanword, and the formation of vowel hiatuses and diphthong alternation patterns predates the borrowing of *sais into Seediq, thus resulting in the suffixed forms of *sais not undergoing the same changes. Proto-Seediq *sais could have been borrowed from Pazih, as Pazih has the exact same form, *sais* 'to sew' (adapted from Li and Tshuchida 2001).

4.4.1.4 Vowel coalescence in Proto-Seediq

Vowel coalescence is a relatively minor change, occurring in environments as described in Section **??**, where the stem must have the structure CVVCV(C). When such a stem is combined with a suffix, the original VV sequence in the antepenultimate and penultimate positions merges into a neutral vowel. Relevant examples are shown in Table **??**, such as *pələalay ~ *pələlayun 'to do first' and *kəicug ~ *kəcugun 'to fear'.

Table 4.39: *u~*ə alternation in Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
*V.V ~*ə	*pələalay	*pələlayun	'to do first'
	*kəicug	*kəcugun	'to fear'

4.4.1.5 Apharesis in Proto-Seediq

Apheresis is also evident in Proto-Seediq. When a stem with the structure VCV(C) is combined with a suffix, the first V is deleted due to the lack of onset protection. Relevant examples are shown in *imah ~ *mahun 'to drink' and *apa ~ *paun 'to carry on back' in Table **??**. In the table, the deleted vowels are represented by an underscore

“_” .

Table 4.40: Apharesis in Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
V~∅	***i**mah	*_mahun	‘to drink’
	***a**pa	*_paun	‘to carry on back’

4.4.2 Consonant alternations in Proto-Seediq

4.4.2.1 Alternations between *-t and *-d- in Proto-Seediq

As mentioned in Section **??**, Proto-Seediq does not allow voiced stops to occur in word-final positions. Therefore, voiced stops like *b and *d undergo devoicing. Additionally, *b is subject to restrictions on word-final labials, which will be discussed later in Section **??**. The alternation between *-t and *-d- is due to this final devoicing. Relevant examples can be found in *həmaŋu**t** ~ *həŋə**d**un ‘to cook’, *məhaga**t** ~ *həga**d**un ‘to pile stones’ in Table **??**.

Table 4.41: Alternations between *-t and *-d- in Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
-t~-d-	*həmaŋu**t**	*həŋə**d**un	‘to cook’
	*məhaga**t**	*həga**d**un	‘to pile stones’

4.4.2.2 Alternations between velar and bilabial consonants in Proto-Seediq

Bilabial consonants including *p, *b, and *m do not occur in the word-final position in Proto-Seediq. However, if we consider the suffixed forms in Proto-Seediq, the alternation between bilabial velar and bilabial sounds can be seen, as shown *madu**k** ~ *du**p**un ‘to hunt (by chasing)’, *məlu**k** ~ *lə**b**an ‘to close’, and *migi**ŋ** ~ *gi**m**an ‘to look for’ in Table **??**.

Table 4.42: Velar and bilabial consonantal alternations in Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
-k~-p-	*madu**k**	*du**p**un	‘to hunt (by chasing)’
-k~-b-	*məlu**k**	*lə**b**an	‘to close’
-ŋ~-m-	*migi**ŋ**	*gi**m**an	‘to look for’

4.4.2.3 Alternation between *-aw and *-ag- in Proto-Seediq

The *-ag sequence cannot occur in the word-final position in Proto-Seediq due to the merger of **-ag and **-aw from pre-Proto-Seediq to Proto-Seediq (see Section **??**, **??** for the reconstruction of *g and *aw). However, when the underlying *-ag- sequence is not in word-final position, i.e., after suffixation, it is not subject to the restriction of word-final position and is realized as *-ag- in the surface. Relevant examples are shown in *rəməŋ**aw** ~ *rəŋ**ag**un ‘to speak’ and *dəmay**aw** ~ *dəy**ag**un ‘to help’ in Table **??**.

Table 4.43: *-aw and *-ag- alternations in paradigm forms of Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
-aw~-ag-	*rəməŋ**aw**	*rəŋ**ag**un	‘to speak’
	*dəmay**aw**	*dəy**ag**un	‘to help’

4.4.2.4 Alternations between *p/*b- and *m- in Proto-Seediq

Alternations between *p/*b- and *m- in Proto-Seediq occur in environments identical to all modern dialects, as mentioned in Section **??**. When the actor voice infix *<əm> is inserted to stems with structures such as *pVCV(C) or *bVCV(C), the dissimilation of two consecutive labial consonants results in the deletion of the initial syllable *pə- or *bə-. Relevant examples are shown in *p**a**tas ~ *m**a**tas ‘to write’ and *b**a**hu ~ *m**a**hu ‘to wash clothes’ in Table **??**.

Table 4.44: *p/*b- and *m- alternations in Proto-Seediq

Alternation	non-suffixed form	suffixed form	Gloss
*p-~*m-	*p**a**tas	*m**a**tas	‘to write’
*b-~*m-	*b**a**hu	*m**a**hu	‘to wash clothes’

4.5 Interim summary

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames

ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tinci-

dunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

References

Anttila, Raimo. 1972. *An introduction to historical and comparative linguistics*. Macmillan New York.

Asai, Erin. 1953. *The Sedik language of Formosa*. Cercle Linguistique de Kanazawa, Kanazawa University.

Blust, Robert. 1999. Subgrouping, Circularity and Extinction: Some issues in Austronesian comparative linguistics. *Selected papers from the Eighth International Conference on Austronesian Linguistics*. Ed. by Elizabeth Zeitoun, and Paul Jen-kuei Li. 1. Symposium Series of the Institute of Linguistics. Taipei, Taiwan: Academia Sinica: 1–54.

Blust, Robert, Stephen Trussel, and Alexander D. Smith. 2023. CLDF dataset derived from Blust's "Austronesian Comparative Dictionary" (v1.2). [Data set]. DOI: https://doi.org/10.5281/zenodo.7741197. URL: https://acd.clld.org/.

Brugmann, Karl, and Hermann Osthoff. 1878. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*. Vol. 1. Hirzel.

Bullock, Thomas Lowden. 1874. Formosan dialects and their connection with the Malay. *China Review* 3.1: 38–45.

Campbell, Lyle, and Mauricio J. Mixco. 2007. *A Glossary of Historical Linguistics*. Glossaries in Linguistics. Edinburgh University Press. ISBN: 9780748623785.

Cheng, Chung-Hua, Dakis Pawan, and Pidu Nokan. 2019. *Seediq Tgdaya Pssgan Ptasan Ngayan : A Handbook of Classified Vocabulary of Seediq Tgdaya Dialect*. Pintung: Hsiang Yuan. ISBN: 9789869715683.

Cheng, Chung-Hua, Nanu Ruwiq, and Masaw Tadaw. 2017. *Patas Bgurah Kari Truku : A Handbook of Classified Vocabulary of Truku*. Pintung: Hsiang Yuan. ISBN: 9789869512930.

Chien, Yuehchen (簡月眞), and Shinji Sanada (眞田信治). 2010. Dong Taiwan Taiyazu de Yilan Keliaoer 東台灣泰雅族的宜蘭克里奧爾 [Yilan Creole of the Atayal People in Eastern Taiwan]. *Taiwan Journal of Indigenous Studies* 3.3: 75–89. DOI: 10.29910/TJIS.201009.0003.

Crowley, Terry, and Claire Bowern. 2010. *An Introduction to Historical Linguistics*. Oxford University Press. ISBN: 9780195365542.

Dyen, Isidore, and Shigeru Tsuchida. 1987. Proto-Austronesian D_5? *A World of Language: Papers Presented to Professor S.A. Wurm on His 65th Birthday*. Dept. of Linguistics, Research School of Pacific Studies, Australian: 173–179.

Fox, Anthony. 1995. *Linguistic reconstruction: an introduction to theory and method*. Oxford University Press, USA.

Goderich, Andre (郭育賢). 2020. *Atayal phonology, reconstruction, and subgrouping*. PhD thesis. Institute of Linguistics, National Tsing Hua University.

Greenberg, Joseph H. 1957. The problem of linguistic subgroupings. *Essays in linguistics*: 46–55.

Hockett, Charles F. 1958. A course in modern linguistics.

Holmer, Arthur J. 1996. *A parametric grammar of Seediq*. Vol. 30. Travaux de l'Institut de Linguistique de Lund. Lund: Lund University Press.

Iwan Perin (依婉 • 貝林). 2005. *Utux, kōngjiān, jìyì yǔ bùluò jiàngòu — yǐ alang Tongan yǔ alang Sipo wéizhǔ de tǎolùn Utux*、空間、記憶與部落建構 — 以 *alang Tongan* 與 *alang Sipo* 爲主的討論 *[Utux, space, memory and tribal construction — discussion focusing on alang Tongan and alang Sipo]*. MA thesis. Department of Indigenous Affairs and Development, National Dong Hwa University. [in Chinese]

Lee, Amy Pei-jung. 2012. The segment /w/ and contrastive hierarchy in Paiwan and Seediq. *Concentric: Studies in Linguistics* 38.1: 1–37.

Lee, Amy Pei-Jung. 2015. Contact-induced sub-dialects in Toda Seediq*. *Asia-Pacific Linguistics (Papers from 12-ICAL, Volume 3, 2015)*: 93–107.

Lee, Amy Pei-jung (李佩容). 2018. *Tàilǔgéyǔ yǔfǎ gàilùn* 太魯閣語語法概論 *[A Grammar of the Truku Language]*. Council of Indigenous Peoples. [in Chinese]

Li, Paul Jen-kuei. 1977. Morphophonemic alternations in Formosan languages. *Bulletin of the Institute of History and Philology, Academia Sinica* 48.3: 375–413.

Li, Paul Jen-kuei. 1981. Reconstruction of proto-Atayalic phonology. *Bulletin of the Institute of History and Philology, Academia Sinica* 52.2: 235–301.

Li, Paul Jen-kuei. 1982. Linguistic variations of different age groups in the Atayalic dialects. *Tsing Hua Journal of Chinese Studies* 14.1–2: 167–191.

Li, Paul Jen-kuei, and Shigeru Tsuchida. 2001. *Pazih dictionary*. Vol. 2. Language and Linguistics Monograph Series. Taipei: Institute of Linguistics, Academia Sinica.

Liao, Shou-chen (廖守臣). 1977. Tàiyǎzú dōng Sàidékèqún de bùluòqiānxǐ (shàng) 泰雅族東賽德克群的部落遷徙（上）[The Tribal Migration and Distribution of the Atayal East Seediq Group (Part I)]. *Bulletin of the Institute of Ethnology, Academia Sinica* 44: 61–206. [in Chinese]

Liao, Shou-chen (廖守臣). 1978. Tàiyǎzú dōng Sàidékèqún de bùluòqiānxǐ (xià) 泰雅族東賽德克群的部落遷徙（下）[The Tribal Migration and Distribution of the Atayal East Seediq Group (Part II)]. *Bulletin of the Institute of Ethnology, Academia Sinica* 45: 81–212. [in Chinese]

Ochiai, Izumi. 2016. Bu-hwan vocabulary recorded in 1874: Comparison with Seediq dialects. *Asian and African languages and linguistics* 10: 287–324.

Ogawa, Naoyoshi. 2006. *A Comparative Vocabulary of Formosan Languages and Dialects*. Ed. by Paul Jen-kuei Li, and Masayuki Toyoshima. Asian and African Lexicon Series 49. Tokyo: Research Institute for Languages, Cultures of Asia, and Africa, Tokyo University of Foreign Studies [《台灣蕃語蒐錄》].

Ogawa, Naoyoshi, and Erin Asai. 1935. *Gengo ni yoru Taiwan Takasagozoku Densetsushū* 原語による臺灣高砂族傳說集 *[Legends of Taiwanese Aboriginal Tribes in their Original Languages]*. Taihoku: Institute of Linguistics, Taihoku Imperial University. [in Japanese]

Ross, Malcolm. 2009. Proto Austronesian verbal morphology: A reappraisal. *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*. Ed. by Alexander Adelaar, and ANdrew Pawley. Pacific Linguistics 601. Canberra: The Australian National University: 295–326.

Sapir, Edward. 1921. *Language: An introduction to the study of speech*. Courier Corporation.

Song, Walis Hian-chi (宋硯之). 2022. Zàilùn yuánshǐtàiyǎ-sàidékèyǔ (Proto-Atayalic) zhōng de *-d 再論原始泰雅-賽德克語（Proto-Atayalic）中的 *-d (Revisiting Proto-Atayalic Final *-d). *Paper presented at the 6th Undergraduate Academic Con-*

ference on Humanities. The Education University of Hong Kong, Tai Po, Hong Kong. [in Chinese]

Song, Walis Hian-chi (宋硯之). 2023. Deconstructing Proto-Atayalic *gʼ. *Paper presented at the 32st Annual Meeting of the Southeast Asian Linguistics Society (SEALS 32)*. Chiang Mai University, Chiang Mai, Thailand.

Song, Walis Hian-chi (宋硯之). 2024a. The Mystery of Atayalic /x/: An Exploration of Its Origin. *Paper presented at the 33st Annual Meeting of the Southeast Asian Linguistics Society (SEALS 33)*. National Tsinghua University, Hsinchu, Taiwan.

Song, Walis Hian-chi (宋硯之). 2024b. Proto-Atayalic *g Revisited: A Mess. *Paper presented at the the 16th International Conference on Austronesian Linguistics (16-ICAL)*. De La Salle University, Manila, The Philippines.

Song, Walis Hian-chi (宋硯之). 2024c. Sàidékèyǔ de lìshǐ céngcì wèntí: Cóng yuánshǐnándǎoyǔ *q de fǎnyìng chūfā 賽德克語的歷史層次問題：從原始南島語 *q 的反映出發 [Historical Strata in Seediq — Examined Through the Reflexes of Proto-Austronesian *q]. *Paper presented at the 25th National Conference on Linguistics (NCL 2024)*. National Tsinghua University, Hsinchu, Taiwan. [in Chinese]

Song, Walis Hian-chi (宋 硯 之), and Samuel Yu-hsiang (潘 俞 翔) Pan. 2024. Morphophonological alternations of stems with final /g/ in Tgdaya Seediq: An OT account. *Paper presented at the 16th International Conference on Austronesian Linguistics (16-ICAL)*. De La Salle University, Manila, The Philippines.

Sung, Li-may (宋麗梅). 2018. *Sàidékèyǔ yǔfǎ gàilùn 賽德克語語法概論 [A Grammar of the Seediq Language]*. Council of Indigenous Peoples. [in Chinese]

Tashiro, Antei (田代安定). 1900. *Taitō Shokuminchi Yosatsu Junkai Nisshi 臺東殖民地豫察巡回日誌 [Taitung Colony Reconnaissance Tour Diary]*. [in Japanese]

Teng, Hsiang-yang (鄧相陽), and Shih-lang Jian (簡史朗). 2023. *Wùshè shìjiàn diàochá yánjiū* 霧社事件調查研究 *[Investigation and Study on Musha Incident]*. Nantou, Taiwan: Taiwan Historica. ISBN: 9789865435837. DOI: https://doi.org/10.978.98654 35/837. [in Chinese]

Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic phonology*. SLCAA Monograph Series 5. Tokyo: Institute for the Study of Languages, Cultures of Asia, and Africa.

Utsurikawa, Nenozō (移川子之藏), Tōichi Mabuchi (馬淵東一), and Nobuto Miyamoto (宮本延人). 1935. *Taiwan Takasagozoku keitō shozoku no kenkyū daiichi satsu* 臺灣高砂族系統所屬の研究 第一冊 *[The Formosan Native Tribes Agen-eralogical and Classificatory Study, Book 1]*. Taihoku: Taihoku Imperial University. [in Japanese]

Yang, Hsiu-fang (楊秀芳). 1976. Sàidéyǔ wùshèfāngyán de yīnyùn jiégòu 賽德語霧社方言的音韻結構 [Phonological structure of the Wushe dialect of Seediq]. *Bulletin of the Institute of History and Philology* 47.4: 611–706.

Yap, Ko-Hua (葉高華). 2011. Dìtú・bālái 地圖・巴萊 [Map Bale]. URL: https://mapsta lk.blogspot.com/2011/11/blog-post_21.html (visited on 01/30/2024). [in Chinese]

Yap, Ko-Hua (葉高華). 2023. *Qiángzhì yízhù: Táiwān gāoshān yuánzhùmín de fēn yǔ lí* 強制移住：臺灣高山原住民的分與離 *[Forced Relocation: The Separation of Taiwan Mountain Indigenous Peoples]*. Taipei: National Taiwan University. ISBN: 9789863507727. [in Chinese]

Yap, Ko-Hua (葉高華). Ongoing. Gāoshān yuán shè wèizhì tú (1931-1945) 高山原社位置圖（1931-1945）[Map of Indiginous Tribes in the Mountainous Area (1931-1945)]. URL: https://www.google.com/maps/d/u/0/viewer?mid=1QIfM5WBOY6s26 bE6NNuIlkpKU-c (visited on 01/30/2024). [in Chinese]