

A network analysis approach to evaluating community language vitality

Jeff ZAHIR

Independent Researcher

Zalmai ZAHIR

University of Oregon

EDITORS:

Gabriela PÉREZ BÁEZ (University of Oregon)

Justin SPENCE (University of California, Davis)

ABSTRACT

Growing efforts to revitalize dormant or near dormant languages have resulted in allocating whatever limited resources to solutions intended to expand or at least maintain the use of these endangered languages. Many involved in these efforts are concerned with planning and creating effective strategies to prevent a language from showing signs of decline towards dormancy. This in turn has generated a need to create evaluative models that provide accurate indicators for the status and health of these languages. What has developed are several sociolinguistic models that focus on measuring a disruption in language use or on measuring endangerment rather than measuring an increase in vitality. This makes it difficult to obtain data for languages that are becoming healthier after a period of decline or even dormancy.

This paper posits a new approach to modeling language community vitality using social network analysis. Network analysis includes mathematical equations to obtain statistics that provide reliable and valid information on social characteristics of a language community. Network analysis statistics can be used to describe a language community's vitality, as well as provide heuristic indicators to assist language revitalizationists to focus their energies on efforts that will help sustain and expand language use.

RESUMEN

Un incremento en los esfuerzos por revitalizar las lenguas dormidas o casi dormidas ha resultado en la asignación de recursos limitados hacia soluciones para expandir o al menos mantener el uso de estas lenguas en riesgo. Muchos de los involucrados en estos esfuerzos se ocupan de planear y crear estrategias efectivas para evitar que una lengua

muestre señales de acercarse a un estado de latencia. Esto ha generado una necesidad de crear modelos de evaluación que proporcionen indicadores hacia el estado y salud de estas lenguas. Esto a su vez ha resultado en el desarrollo de varios modelos sociolingüísticos con un enfoque en medir interrupción en el uso de la lengua o en medir el estado de riesgo en vez de sistemas para medir vitalidad. Esto causa dificultades para obtener datos para lenguas que se encuentran recuperándose de un periodo de desuso o incluso de latencia.

Este artículo propone una nueva manera de abordar modelos para evaluar la vitalidad de una comunidad lingüística usando un análisis de redes sociales. Este tipo de análisis incluye ecuaciones matemáticas para obtener estadísticas fiables sobre las características sociales de una comunidad lingüística. Estas estadísticas pueden usarse para describir la vitalidad de una comunidad lingüística y también para proveer indicadores heurísticos para asistir a los practicantes de la revitalización en enfocar sus energías en aquellos esfuerzos que ayudarán a mantener y expandir el uso de una lengua.

1. INTRODUCTION

There is an evolving movement to revitalize dormant or near dormant languages. Dormant languages are those that no longer have a language community, and no one is putting forth energy to use the language. Growing efforts to revitalize these languages has resulted in allocating limited resources to solutions intended to expand or, at least, maintain the use of these endangered languages (Fishman, 1991). Many involved in these efforts are concerned with planning and creating effective strategies for the language of a community to not show signs of decline towards dormancy. This in turn has generated a need to create evaluative models that provide accurate indicators for the status and health of these languages. What has developed are several models that attempt to measure a language community's vitality.

The term *vitality* refers to the energy put forth by a language community to communicate using the target language. Sociolinguistic models of language vitality assessment focus on measuring language use disruption or endangerment rather than increase in vitality (Fishman, 1991; UNESCO Ad Hoc Group on Endangered Languages, 2003). The focus on disruption and endangerment makes it difficult to obtain data for languages on a trajectory to becoming healthier after a period of decline or even dormancy. As such, they operate as deficit-based models and are not adequate for measuring a reversal of language shift and increased vitality. These types of models also use vague terminology such as 'speaker' and 'community' that must be defined by the user, making such models difficult to apply (see Section 2.1 below).

This paper posits a new approach to modeling language community vitality using social network analysis, a well-established method for modeling social network behavior and the connections created and maintained between people. It includes mathematical equations to obtain statistics that provide reliable and valid information on social characteristics of a community, including speaking communities. In combination with data gathered over time (longitudinal data), network analysis statistics can be used to describe a language community's vitality, as well as provide heuristic (investigative) indicators to assist language revitalizationists to focus their energies on efforts that will help sustain and expand language use.

The use of network analysis to analyze language community vitality is motivated by and based on preliminary efforts to model language use in communication of several awakening languages. *Awakening languages* are those that were once dormant or near dormant and are being given new energy by a language community to be used for communication. Like dormant plants in the winter that reemerge in the spring, languages can reemerge when people put forth renewed energy into their use. Methods used for awakening languages include home-based language nests in concert with a conversation component and other language socialization projects. *Language nests* are physical locations where the language is spoken, thus, a *home-based language nest* is a nest located within a home where only the target language is used (Zahir, 2018). Some of the languages for which this approach is taken include Lushootseed (lut) (Zahir, 2018), mitsqanaqan (veo), Nuu-wee-ya' (ctc/tol) (Viles, this special issue), Western Mono (mnr), and Pit River (acv). They had all experienced decades of language decline and even dormancy and are now experiencing an exciting increase in language community vitality. In these language communities, home-based language nesting results in individuals developing varying degrees of proficiency. There are those who speak their language comfortably in a variety of domains within their language nest, as well as family members and friends who can somewhat participate with these speakers, but they have yet to develop the skills to fulfill a greater range of communicative needs. Proficiency analysis approaches are difficult to apply to communities like Lushootseed, where the complexity in proficiency evaluation increases when the language community population surges to several hundred. Sociolinguistic deficit-based models have aspects that do not apply to language communities rebounding from (near) dormant situations. To remedy this, the social network model we propose veers away from assessing speaker proficiency and language deficiency to focus on social language interaction by members of a language community.

In writing this paper, we have in mind the language revitalization practitioner within a community who may want to implement the methods we propose to evaluate the growing vitality of their language. Therefore, we write with a language practitioner audience in mind. In Section 2, we provide an overview of the literature on language endangerment and vitality to explain how past approaches fail to adequately recognize gains in vitality after a period of (near) dormancy. Then, in Section 3, we

introduce social network theory and discuss, in accessible terms, key concepts and the reasons we are proposing a social network approach to assessing language vitality growth. In Section 4, we provide a case study and walk the reader through a social network analysis of the gains in vitality of the mitsqanaqan language. Section 5 closes with a brief conclusion.

2. MEASURING LANGUAGE VITALITY

Linguists generally agree that the vitality of languages experiencing language shift depends on a complex interplay of identity, community, and polity (Fishman, 1991; Giles et al., 1977; Lewis & Simons, n.d.; Rehg & Campbell, 2018), thus, it is in revealing the complex relationships between these social factors that makes measuring language vitality so difficult and so important. Since the early 1960's linguists, sociologists, and psychologists have been trying to tease out these relationships with varying degrees of success, using proxy measures of language use such as the number, percent, and fluency of speakers. The social sciences are full of examples of this type of problem. Analysts will often create an index of multiple measurements whenever confronted by a hidden mechanism of complex variables that explain a problem (e.g., Consumer Price, Social Progress, or student GPA). The requirements for factors used in any index are that they be structurally linked in some consistent way and that they can be normalized in such a way that no single factor outweighs another. One such index is the Language Endangerment Index and related indicators (Fishman, 1965, 1991; Krauss, 2007; Lee & Van Way, 2016; Lewis & Simons, 2010) which provide linguists with an assessment tool for measuring the degree of a language's vitality or endangerment. These indices build a composite using 8 or 9 factors that result in a "level" designation from one to ten indicating a language's overall vitality. Each varies in interpretation of the factors, but all come with caveats, weighting factors and data collection methods. The authors emphasize that these methods are most useful with large datasets and only work in the aggregate.

In cases where large datasets are unobtainable, causal relationships and the mechanisms of language endangerment can be even more elusive. What is needed is a dataset that builds on the relationships of speakers and the factors that motivate their communications. Milroy & Milroy (1985) suggests that factors affecting the density of a network of speakers in a community are indicators of "ethnolinguistic strength." Several other specialists in sociolinguistics use a similar heuristic to test known factors (Giles et al., 1977). Once unambiguous characteristics of the speakers (new, existing, former, and non-) are defined, similar clarity is required for defining the actions that connect them. For language speakers, the simplest connection is that they speak to one another, but the analyst may wish to add layers of complexity, such as the amount of time spent speaking or measures of proficiency.

2.1 AN EXERCISE IN USING UNESCO'S LANGUAGE VITALITY AND ENDANGERMENT FRAMEWORK WITH MITSQANAQAN

To exemplify the need for a new approach to modeling language vitality for an emerging near dormant or dormant language, below we attempt to assess the mitsqanaqan language using UNESCO's sociolinguistic model (UNESCO Ad Hoc Group on Endangered Languages, 2003), a model designed with 9 factors that work together to assess the vitality of a language. 8 of the 9 factors have a grading between 0 [low] and 5 [high]. Grade 0 means the language is 'extinct' and Grade 5 means the language is at the highest level of vitality. These 9 factors and grades are then put into a table as a framework to give an assessment of language health.

FACTOR 1: INTERGENERATIONAL TRANSMISSION

Factor 1 assesses the intergenerational transmission of a language. It assumes that a healthy language is spoken across all generations between the eldest age group down to a younger group (Grade 5, Table 1), and a critically endangered language is only used by great great-grandparents (Grade 1, Table 1) with Grades 4, 3 and 2 applying to languages with a gradually older speaking population. Factor 1 assumes an initially healthy language dwindling in number of speakers, beginning with the youngest generation because children are not learning their language from their parents. This paradigm does not apply to mitsqanaqan because the language went dormant 50 years ago when the eldest speakers passed. Today, the mitsqanaqan language community is mostly adults within the 20-to-40-year age range, while the eldest members of the Ventureño Chumash tribe do not speak it. Therefore, grading the mitsqanaqan language from 5 to 0 in Table 1 does not apply and cannot be used as an accurate assessment of mitsqanaqan's vitality.

Grade	Degree of endangerment	Speaker Population
Grade 5:	Safe	The language is used by all ages, from children on up
Grade 4:	Unsafe	Most children are speaking
Grade 3:	Definitely endangered	Most language used is by an older generation
Grade 2:	Severely endangered	Most language used is by grandparents and up
Grade 1:	Critically endangered	Only great great-grandparents speak
Grade 0:	Extinct	No speakers exist

Table 1: Grading for Factor 1

FACTOR 2: ABSOLUTE NUMBER OF SPEAKERS

Factor 2 assesses the absolute number of speakers of a language. The term 'speaker' is not defined, so it is difficult to assess who should be included in this factor. For the purposes of this

exercise, we will not try to define speaker but rather use the reported number of members of the language community from our survey (mentioned above). For mitsqanaqan, this number is 25. Factor 2 is the only factor that does not have a grading format.

FACTOR 3: PROPORTION OF SPEAKERS WITHIN THE TOTAL POPULATION

Factor 3 compares a speaking population within a total population of a community (Table 2). This factor is difficult to assess given that very few communities are isolated. Therefore, how can a total population be defined? For mitsqanaqan, if it is only the total population of the tribally enrolled Ventureño Chumash community, then members who are not enrolled would be excluded, which would reflect an inaccurate assessment of the mitsqanaqan community because there are members of the language community who are not enrolled in the tribe. Because a community that accurately represents the speaking population cannot be defined, Factor 3 cannot be used.

Grade	Degree of endangerment	Portion of speakers within the total reference population
Grade 5:	Safe	All speak the language.
Grade 4:	Unsafe	Nearly all speak the language.
Grade 3:	Definitely endangered	A majority speak the language.
Grade 2:	Severely endangered	A minority speak the language.
Grade 1:	Critically endangered	Very few speak the language.
Grade 0:	Extinct	None speak the language.

Table 2: Grading for Factor 3

FACTOR 4: TRENDS IN EXISTING LANGUAGE DOMAINS

Factor 4 addresses the social interaction of a language in terms of function and domain (Table 3). The terms for language ‘domain’ and ‘function’ are ambiguous, and the term ‘dwindling domains’ for Grade 3 is deficit-focused. This grade further emphasizes a deficit-based approach by describing Grade 3, as in the case where “... the dominant language begins to penetrate even home domains” (p.10). This deficit does not work for languages being revitalized. In these cases, the target language is not dwindling for home domains, but rather increasing and the dominant language is beginning to decrease. However, an intermittent grade between Grades 2 and 4 is needed to describe what is occurring with mitsqanaqan. This language is occurring both within and outside the home. Within the homes there are language nests and communication between family and friends. Outside of the home, language occurs at some tribal gatherings and some social interaction between speakers. If Grade 3

is redefined as a middle level between Grades 2 and 4, then the mitsqanaqan language community receives a Grade 3 (Grade 3 is in bold for clarity).

Grade	Degree of endangerment	Domains and Functions
Grade 5:	Universal use	The language is used for all domains and for all functions.
Grade 4:	Multilingual parity	Two or more languages may be used in most social domains and for most functions.
Grade 3:	Redefined Dwindling domains	A grade between 4 and 2. The language is in home domains and for many functions, but the dominant language begins to penetrate even home domains.
Grade 2:	Limited or formal domains	The language is used for limited social domains and several functions.
Grade 1:	Highly limited domains	The language is used only in restricted domains and for a very few functions.
Grade 0:	Extinct	The language is not used for any domain nor for any function.

Table 3: Grading for Factor 4

FACTOR 5: RESPONSE TO NEW DOMAINS AND MEDIA

For response to new domains and media, the word ‘domain’ is not defined and in addition, the term ‘new domain’ has ambiguity. However, mitsqanaqan is evolving vocabulary for objects in today’s typical home, for example, words for ‘refrigerator’, ‘toothbrush’, ‘bathroom’ and ‘sink’. In addition, new language apps and affordances to write mitsqanaqan on the Internet have expanded its use into social media. Therefore, mitsqanaqan gets a grade of 2 (See Table 4).

Grade 5:	Dynamic	The language is used in all new domains.
Grade 4:	Robust/active	The language is used in most new domains.
Grade 3:	Receptive	The language is used in many domains.
Grade 2:	Coping	The language is used in some new domains.
Grade 1:	Minimal	The language is used only in a few new domains.
Grade 0:	Inactive	The language is not used in any new domains.

Table 4: Grading for factor 5

FACTOR 6: MATERIALS FOR LANGUAGE EDUCATION AND LITERACY

mitsqanaqan has a developing dictionary, a grammatical sketch and pedagogical materials, but the majority of resources are in archival form with the greatest contribution from J. P. Harrington. Therefore, the mitsqanaqan language has a grade of 2 for Factor 6 (see Table 5).

Grade 5:	There is an established orthography, literacy tradition with grammars, dictionaries, texts, literature, and everyday media. Writing in the language is used in administration and education.
Grade 4:	Written materials exist, and at school, children are developing literacy in the language. Writing in the language is not used in administration.
Grade 3:	Written materials exist and children may be exposed to the written form at school. Literacy is not promoted through print media.
Grade 2:	Written materials exist, but they may only be useful for some members of the community; and for others, they may have a symbolic significance. Literacy education in the language is not a part of the school curriculum.
Grade 1:	A practical orthography is known to the community and some material is being written.
Grade 0:	No orthography is available to the community.

Table 5: Grading for Factor 6

Grade 5:	Equal support	All languages are protected.
Grade 4:	Differentiated support	Minority languages are protected primarily as the language of the private domains. The use of the language is prestigious.
Grade 3:	Passive assimilation	No explicit policy exists for minority languages; the dominant language prevails in the public domain.
Grade 2:	Active assimilation	Government encourages assimilation to the dominant language. There is no protection for minority languages.
Grade 1:	Forced assimilation	The dominant language is the sole official language, while non-dominant languages are neither recognized nor protected.
Grade 0:	Prohibition	Minority languages are prohibited.

Table 6: Grading for Factor 7

FACTOR 7: GOVERNMENTAL AND INSTITUTIONAL LANGUAGE ATTITUDES AND POLICIES, INCLUDING OFFICIAL STATUS AND USE

The mitsqanaqan language is moderately well supported by tribal, local, state and Federal Government and has a Grade 4 for Factor 7 (see Table 6).

FACTOR 8: COMMUNITY MEMBERS' ATTITUDES TOWARDS THEIR OWN LANGUAGE

The mitsqanaqan language is moderately well supported by community members and has a Grade 4 for Factor 8 (see Table 7).

Grade 5:	All members value their language and wish to see it promoted.
Grade 4:	Most members support language maintenance.
Grade 3:	Many members support language maintenance; others are indifferent or may even support language loss.
Grade 2:	Some members support language maintenance; others are indifferent or may even support language loss.
Grade 1:	Only a few members support language maintenance; others are indifferent or may even support language loss.
Grade 0:	No one cares if the language is lost; all prefer to use a dominant language.

Table 7: Grading for Factor 8

FACTOR 9: AMOUNT AND QUALITY OF DOCUMENTATION

As mentioned under Factor 6, mitsqanaqan has some documentation materials but most are in archive format. Work is being done to digitize archival records to make them available to the tribal and language community, however, very little audio was recorded of the mitsqanaqan 1st-language speakers. Therefore, mitsqanaqan receives a grade 4 for Factor 9 (see Table 8).

Grade 5:	Superlative	There are comprehensive grammars and dictionaries, extensive texts; constant flow of language materials. Abundant annotated high-quality audio and video recordings exist.
Grade 4:	Good	There is one good grammar and a number of adequate grammars, dictionaries, texts, literature, and occasionally updated everyday media; adequate annotated high-quality audio and video recordings.
Grade 3:	Fair	There may be an adequate grammar or sufficient amount of grammars, dictionaries, and texts, but no everyday media; audio and video recordings may exist in varying quality or degree of annotation.

Grade 2:	Fragmentary	There are some grammatical sketches, word-lists, and texts useful for limited linguistic research but with inadequate coverage. Audio and video recordings may exist in varying quality, with or without any annotation.
Grade 1:	Inadequate	Only a few grammatical sketches, short word-lists, and fragmentary texts. Audio and video recordings do not exist, are of unusable quality, or are completely un-annotated.
Grade 0:	Undocumented	No materials exist.

Table 8: Grading for Factor 9

In summary, mitsqanaqan can be characterized by the UNESCO Language Vitality and Endangerment framework in Table 9. This assessment model can be used to analyze some of the characteristics of the mitsqanaqan language, however, not without difficulties. Factor 1 attempts to grade intergenerational transmission based upon the age of the youngest generation of speakers. It is not applicable to languages that have lost their eldest speaking group and revitalization efforts are from a younger generation. Factor 2 lists the absolute number of speakers. This is challenging because the term ‘speaker’ is not defined. Users of the UNESCO framework will have to define this term for themselves. The likelihood that languages will evolve different definitions for ‘speaker’ makes any cross linguistic comparison problematic. Factor 3 grades the proportion of speakers within the total population. This factor is difficult to apply until an accurate definition for ‘total population’ is made that is inclusive of all members of a language community. Factor 4 grades trends in existing language domains. It is not clear what language ‘domain’ and ‘function’ are and the term ‘dwindling domains’ does not apply to languages that are expanding. Like Factor 1, this term focuses on a growing deficiency and morbidity of a language. Factor 5 grades the response of a language to new domains and media. Factors 6 through 9 seem to be evaluable without problems. However, these last four factors are subjective in measuring language vitality and/or endangerment.

This assessment model is only partially effective in assessing an awakening language. It uses subjective reasoning and focuses on deficits in the use of the language. Even though the title for this assessment asserts that this model assesses vitality, wording within the factors indicates otherwise. Terms such as ‘dwindling domains’, ‘gradually older speaking population’ and ‘extinct’ languages imply that this model is focused on language morbidity and mortality. This is unfortunate given the move that the field of language revitalization has made for over a decade to not rely on death metaphors to describe language endangerment, much less vitality and revitalization (Leonard, 2011). This is why this exercise in trying to assess mitsqanaqan with this sociolinguistic approach underscores the

urgency in developing objective empirical assessment modeling techniques that language practitioners can use to more accurately understand the vitality status of an awakening language.

Factor 1			Non-applicable
Factor 2			25 speakers
Factor 3			Non-applicable
Factor 4	Grade 3:	Redefined Dwindling domains	A grade between 4 and 2. The language is in home domains and for many functions, but the dominate language begins to penetrate even home domains.
Factor 5	Grade 2:	Coping	The language is used in some new domains.
Factor 6	Grade 2:		Written materials exist, but they may only be useful for some members of the community; and for others, they may have a symbolic significance. Literacy education in the language is not a part of the school curriculum.
Factor 7	Grade 4:	Differentiated support	Minority languages are protected primarily as the language of the private domains. The use of the language is prestigious.
Factor 8	Grade 4:	Good	There is one good grammar and a number of adequate grammars, dictionaries, texts, literature, and occasionally updated everyday media; adequate annotated high-quality audio and video recordings.
Factor 9	Grade 2:	Fragmentary	There are some grammatical sketches, word-lists, and texts useful for limited linguistic research but with inadequate coverage. Audio and video recordings may exist in varying quality, with or without any annotation.

Table 9: Summary of Factors 1 through 9 as applied to mitsqanqan

3. PREVIOUS SOCIAL NETWORK MODELS FOR LANGUAGE

Social network studies have been used in sociolinguistics starting with Labov’s foundational work that uses network modeling to create a sociogram or network visualization of a group of inner-city youth (1972) (Figure 1).

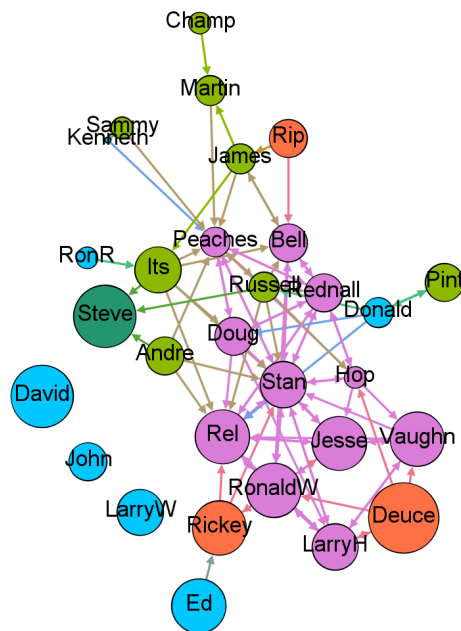


Figure 1: Labov's network graph of inner city youth (Labov, 1972)

Labov's sociogram depicts the degrees of vernacular use by members of a street gang who, in the model, are represented by nodes (or circles). The ties connecting the nodes indicate that those members identify each other as part of the same gang. An analysis of the youths' vernacular language use shows that members with more ties use more of the group's vernacular. Labov argues that people who are more central to a network will likely use more of that network's distinctive vernacular and vice-versa.

Other important contributions include Milroy (1987), who developed a network model for measuring sociolinguistic variation in her study of working-class speech in Belfast. Building upon Labov's work, Milroy expands the social network characteristics that can affect speech behavior, including network zones, density of connectedness, clusters, and multiplexity of relationships (Milroy, 1987). In Figure 2, Milroy uses a network graph to model part of the Clonard community where she gathered dialectal data. Nodes represent individuals and families with whom she spoke and the ties between the nodes show who is talking to whom. In this case, all parties know and speak to each other, meaning the community is 100% connected.

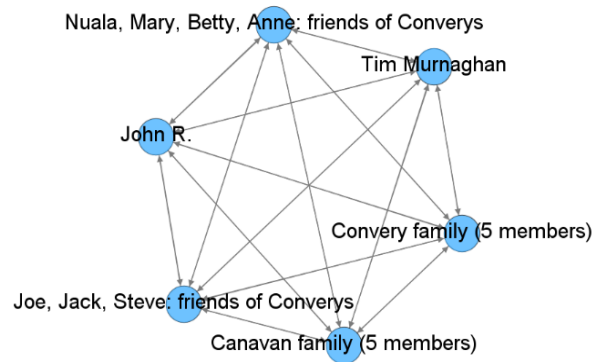


Figure 2: A portion of the Clonard community network (Milroy, 1987)

Although no studies have applied network modeling to language revitalization research, Joshua Fishman's eight stage Graded Intergenerational Disruption Scale (GIDs) (1991) includes social aspects of linguistic behavior to determine whether a language is endangered.¹ Stage 1 represents the full revitalization of a language on a local and national scale and Stage 8 describes the case of a highly endangered language. A brief synopsis of the GIDs grading system follows:

- Stage 8: speakers are socially isolated elders who do not speak to younger generations
- Stage 7: speakers are socially integrated but beyond childbearing age and so cannot directly transmit the language to younger generations
- Stage 6: language is transmitted orally and cross-generationally
- Stage 5: literacy in the home, school, and local community
- Stage 4: language in lower education
- Stage 3: language used in local workplace spheres and can intersect with other communities
- Stage 2: language used in local government, communities, and media
- Stage 1: language in higher education and national government

Stages 8 through 6 focus on language endangerment in terms of degree of socialization between speakers. Communities in which language use occurs among a large swath of the population have more language vitality than in one where only a few isolated speakers remain. Although not explicit, a social aspect of language endangerment can also be garnered from stages 5 through 1 because social interaction is inherent in educational institutions, workplaces, and government.

¹This analysis can be adapted to the Expanded Graded Intergenerational Disruption Scale (EGIDS) proposed in Lewis and Simons (2010).

These three foundational works discuss dynamics of social interactions that influence language behavior, and it is exactly these social network constructs that are foundational in assessing language health. If language is used to convey ideas and information between people (Sapir, 1921), then as revitalizationists we are necessarily concerned with language behavior that cultivates more communication and idea sharing between more individuals. In particular, we want to address the precept that social interaction and group influences, not just individual influences, cause changes in language-use behaviors. Hence, this is why we are positing a paradigm shift from viewing language vitality in terms of deficit-based frameworks that focus on language mortality to one that brings focus to a network approach that measures the linguistic social interactions between people within a language group that lead to expanding language use.

Next, we will begin our presentation on network analysis and language community vitality. We will start with a description of the mitsqanaqan language, from which we will be using data for our discussion. We will then present the basics of social network analysis and use actual mitsqanaqan data to show how to use social network analysis techniques to describe the vitality of a language community.

4. NETWORK ANALYSIS AND LANGUAGE COMMUNITY VITALITY

4.1 MITSQANAQAN LANGUAGE COMMUNITY

To demonstrate the efficacy of social network analysis techniques in assessing vitality, we will turn to the case of data from the mitsqanaqan language community. mitsqanaqan, also known as Ventureño, is part of the Chumashan language family. Its territory is within Southern California where its borders extend north to Ventura, south to Malibu, and include the Simi Hills (“Ventureño Language,” 2022). mitsqanaqan was a dormant language for over 50 years with the last 1st-language speaker dying in the 1960s. Recent efforts to reawaken mitsqanaqan have created an expanding language community working on revitalizing and stabilizing the language. For the past 3 years (since 2020), ongoing classes, which cover how to do self-narration in the language while doing daily activities, have cultivated language socialization between family members and friends. This process is called ‘reclaiming domains’ and is used in home-based language nesting.

As mentioned in the introduction, nests are allocated spaces in the home, such as the kitchen, where only the target language is spoken. Once an individual has reclaimed enough domains in an allocated area, then they can designate that area as a language nest, creating a space where family and friends hear and use language (Zahir, 2018; Viles, this special issue). Instruction on conversation coincides with reclaiming domains and nesting which develops skills for speakers to converse with family and friends. The resulting effect of these approaches is that people are speaking more. People

still learn skills in language comprehension and production, but these abilities are mostly obtained through increasing use and interaction in the language.

To gather data on the mitsqanaqan language community, we conducted an online survey that asked participants to list at least three people to whom they speak and describe the intensity of the relationship. Intensity was graded 1 to 5, where 5 was the highest. For analyzing the mitsqanaqan language community, we used the network analysis software Gephi for network visualization and analysis (*Gephi - The Open Graph Viz Platform*, n.d.).²

4.2 NETWORK GRAPHS

4.2.1 DEFINING LANGUAGE COMMUNITY VITALITY

Before we begin our presentation on social network analysis, it is important to understand that the focus of this chapter is to assess language community vitality, not language revitalization. In general, language revitalization means any vigorous energy put forth to promote language use (Hinton, 2001, p. 5). This includes actions that might not directly cause a change in language behavior but can play important roles in making it possible for language behavior change, especially if the language is dormant. It is not our objective to present an analysis of the efficacies of these activities. Instead, we focus on how to measure changes in a language community in terms of language expansion, contraction, and/or maintenance. Even though this chapter does not focus on the causes of language behavior change, we do refer to and draw from similar mitsqanaqan community data that examines specific language socialization methodologies (briefly mentioned above).

In addition, we forego the need to define the terms ‘speaker’ and ‘fluency’, which can be difficult and inconsistent cross linguistically. Instead, our approach models ‘participants’ of a language community who have speaking relationships with others within the language community, including relationships with those who speak; those who speak and listen; and those who just listen. For example, such relationships can occur within a single household in which parents speak the target language to each other and to their child, but their child responds in another language. Speaking and listening occurs between the parents, but they are only speakers when speaking to the child. Conversely, the child is only a listener until they learn speaking skills and/or choose to respond in the target language. Including those who only listen as part of the ‘language community’ bypasses the need for proficiency to be required for membership and, thus, results in a more robust community. This more inclusive approach does not undermine the logic that good language skills enhance more language use—as we discuss below, a strong language community is conditioned by more people speaking to one another with more intensity.

²Gephi is a widely used and freely available open-source network analysis software.

4.2.2 NETWORK GRAPH ELEMENTS³

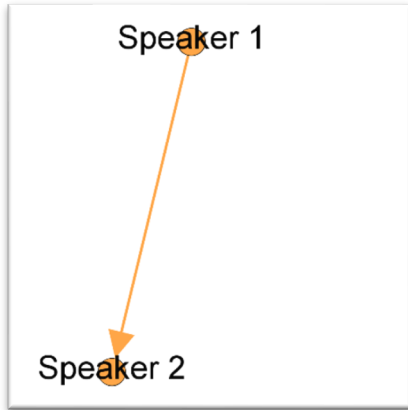
When individuals have social interactions within a language group, then these individuals are members of that language community. A language community is a type of network which is amenable to social network analysis. In turn, social network analysis is the study of graphs that represent relationships between distinct objects or people (Light & Moody, 2020). In particular, language communities are a type of social network, defined as consisting of individuals who have social interactions with one another. There is sizable literature that covers valid and reliable techniques for measuring social networks (Jackson, 2008). These techniques involve using equations and algorithms that work for all types of social networks, and therefore can be used cross linguistically for measuring characteristics of a community regardless of the language's condition.

A network graph of a language community consists of nodes to represent the participants and ties between the nodes to represent who is talking to whom. The four graphs in Figure 3 represent simplified networks consisting of two members of a language community. Arrows at the ends of the ties represent the direction of speech. Ties with arrows are called 'directed ties.' When there are arrows at both ends of a tie, the tie is 'bidirected' and represents both participants speaking to each other (Figure 3c). When there are no arrows, the tie is 'undirected'. Undirected ties mean that there is a social interaction, but the direction of the interaction is unspecified (Figure 3d) (Jackson, 2008). Networks that include directionality are sometimes called digraphs.

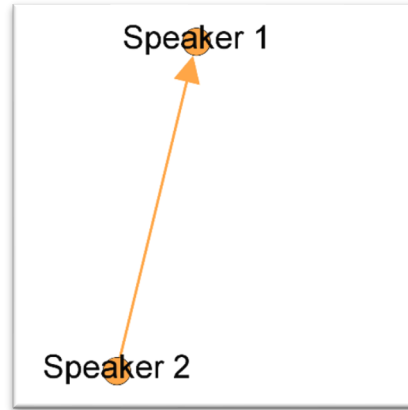
'Degree' is a characteristic of a participant. It is the number of ties a participant has (Jackson, 2008). Figure 4 is a graph of 6 participants. There are 4 bidirected ties between participant 1 and participants 2 through 5; and 1 directed tie from participant 1 to participant 6. Participant 1 has 4 arrows pointing towards it and 5 arrows pointing away from it. The 4 ties pointing towards participant 1 are called the 'in-degree', and the 5 ties pointing away from participant 1 are called the 'out-degree'. The summation of the in-degree plus the out-degree is a participant's 'degree'. For example, the degree of participant 1 is $4+5=9$. Table 10 lists the in-degree, out-degree, and degree for each participant in Figure 4. Participant 1 has the largest degree of 9 and participant 6 has the smallest degree of 1. The sizes of the participants in Figure 4 show this characteristic where participant 1 is the largest circle and participant 6 is the smallest.

³ Note that the precepts of network analysis we present below have previously been covered (Jackson, 2008; Light & Moody, 2020; Newman, 2010).

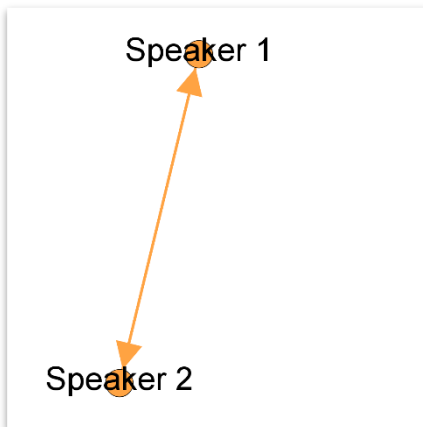
(a) Speaker 1 speaks to 2 (1 tie)



(b) Speaker 2 speaks to 1 (1 tie)



(c) Speakers 1 and 2 speak to each other (2 ties)



(d) non-directional (1 tie)

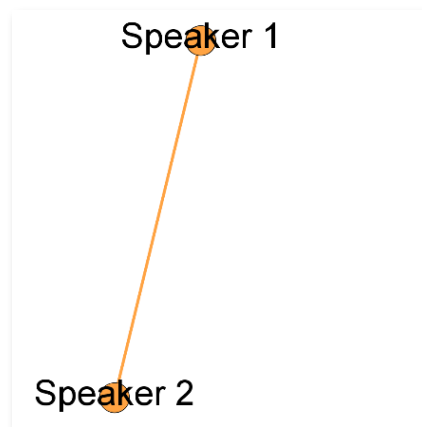


Figure 3: Directed Language Ties

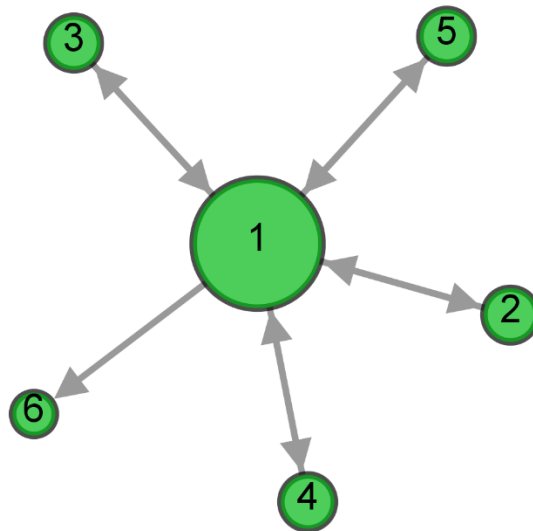


Figure 4: Example of Participant Degree

Participant	Indegree	Outdegree	Degree
1	4	5	9
2	1	1	2
3	1	1	2
4	1	1	2
5	1	1	2
6	1	0	1

Table 10: Degrees of Participants in Figure 4.

Ties can vary by qualitative or quantitative characteristics. A network with valued or weighted ties is called a weighted network. In this case, the value of a tie represents the intensity with which a participant speaks to another participant (Jackson, 2008). In Figure 5, participant 1 speaks to participants 2 and 3. The values for the tie strength are given in Table 11 where the first column titled 'Source' is the speaking participant and the second column titled 'Target' is the listener. The value is 1 for the tie to participant 2 and the value is 2 for the tie to participant 3. This means that, for a range of 1 to 2, the intensity in terms of time participant 1 speaks to participant 3 is twice that of participant 2. In Figure 5, the thicker line and larger arrowhead with participant 3 indicates that the value of the tie is greater than the value of the tie with participant 2.

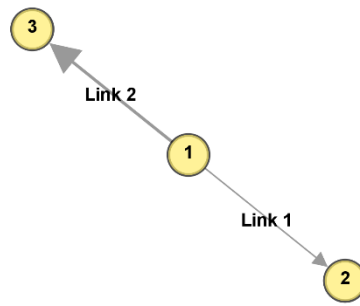


Figure 5: Thickness of Line Indicates Intensity

Source	Target	Value
1	2	1
1	3	2

Table 11: Value of Ties in Figure 5

Where a tie is a link between participants, a series of consecutive ties between participants is called a *pathway*. The shortest pathway between any two participants is called the *direct pathway*. The *diameter* of a language community is the number of ties in the largest direct pathway between any two people within the language community (Jackson, 2008). In Figure 6, participants 2 and 3 are the farthest apart, and the shortest distance between them is 2 ties, giving this language community a diameter of 2. This means that participants 2 and 3 are 2 ties away from speaking to one another.

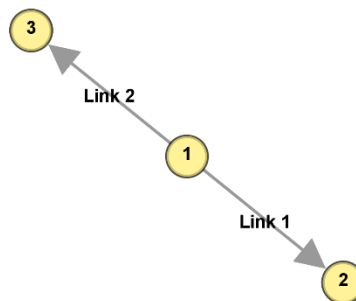


Figure 6: Network of 3 Speakers and a Diameter of 2

These are just a few of the basic elements of social network analysis. Next, we will discuss how these elements can be used to generate statistics that describe characteristics of a social network and in turn, we will show how these characteristics can be used to assess the health of a language community.

4.2.3 NUMBER OF PARTICIPANTS

As previously mentioned, the participants of a language community vary in type. They include people who speak often to several other participants, as well as those who listen but do not speak. The total sum of participants is the total membership of a language community. In terms of vitality, an increase in membership over time indicates language growth; a decrease indicates language contraction; and no change in membership indicates language maintenance.

The mitsqanaqan language community is graphed in Figure 7 below, and the participants are listed in Table 12. There are a total of 25 participants. If language socialization efforts continue, this number should increase over time.

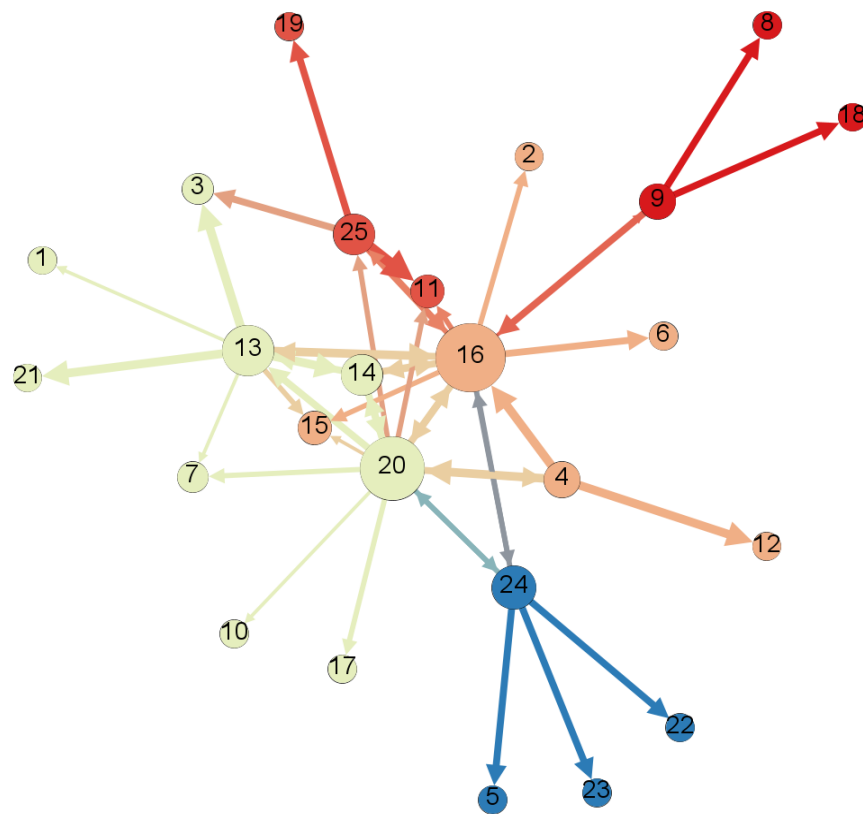


Figure 7: Network Graph: mitsqanaqan Language Community

4.2.4 NUMBER OF TIES

Like participants, ties also vary in type. As mentioned above, ties can be directed, bidirected and non-directed. Ties also vary in intensity where some have high values while others have small values.

Regardless of the type, the number of ties within a language community is an indicator to gauge language community vitality because it represents the number of language-based social interactions between participants. An increase in the number of ties indicates language expansion; a decrease indicates language contraction; and no change indicates language maintenance.

Table 12 lists the ties for the mitsqanaqan language community. Column 1 lists the number of ties followed by the source, target, and value or weight. The total number of ties for the mitsqanaqan language community is 46 (listed in the final row in bold for clarity).

We are not suggesting that a larger number of ties outweighs a smaller language community with fewer ties but has participants who speak with more intensity to each other. Rather, the number of ties is only one characteristic of the language community to use in conjunction with other statistics to understand the language community and its vitality.

Tie number	Source	Target	Value	Tie number	Source	Target	Value
1	4	12	4	27	20	4	3
2	4	16	4	28	20	7	2
3	4	20	4	29	20	10	1
4	9	8	3	30	20	11	2
5	9	16	3	31	20	13	3
6	9	18	3	32	20	14	3
7	13	1	1	33	20	15	1
8	13	3	4	34	20	16	3
9	13	7	1	35	20	17	2
10	13	14	4	36	20	24	2
11	13	15	2	37	20	25	2
12	13	16	4	38	24	5	3
13	13	21	4	39	24	16	2
14	14	13	3	40	24	20	2
15	14	16	3	41	24	22	3
16	14	20	3	42	24	23	3
17	16	2	2	43	25	3	3
18	16	6	3	44	25	11	5
19	16	9	1	45	25	16	3
20	16	11	2	46	25	19	3
21	16	13	3				

22	16	14	3
23	16	15	2
24	16	20	3
25	16	24	2
26	16	25	2
Total = 46			124

Table 12: Mitsqanaqan Community Ties

4.2.5 AVERAGE DEGREE

Where degree is a characteristic of a participant, the average degree throughout the network is a characteristic of the language community. The average degree in a directed network is the sum of the ties divided by the number of participants. Where the number of ties is 46 and the number of participants is 25, the average degree for the mitsqanaqan language community is $46 \div 25 = 1.84$. An increase in the average degree indicates language expansion; a contraction indicates language contraction; and no change indicates language maintenance. As participants speak to more people, this number will increase.

4.2.6 AVERAGE VALUE OF TIES

The average value of the ties is another network statistic useful in the analysis of a language community. The average value is the sum of the values divided by the number of ties. Table 12 above lists the values for the mitsqanaqan language community. The value is on a scale from 1 to 5 where 5 is the highest value of use. The total sum of the values is 124 (listed in the final row of Table 12 in bold for clarity). The total number of ties is 46. Therefore, the average value for the mitsqanaqan language community is $124 \div 46 = 2.70$. As more participants become more adept at speaking and include more ways to use the language, the average value of the language community should increase over time.

4.2.7 DENSITY OF LANGUAGE COMMUNITY

The 'density' of a social network is calculated by dividing the actual number of ties with the total possible ties. The total possible ties of a community: $N \times (N - 1)$ where N = the number of participants. For the mitsqanaqan community, the number of participants, N , is 25 and $25 \times (25 - 1) = 600$ possible ties. Since the actual number of ties is 46, the density of the mitsqanaqan community is $46 \div 600 = 0.077$. As participants speak more to others, the number of ties should increase and, thereby, increase the density of the language community. In terms of language vitality, as density approaches 1, the community is becoming more interconnected, indicating that more people are speaking more language to each other, expanding the language. Conversely, if density decreases, then the language is

contracting; and if density does not change, it is maintaining. If membership of a language community is growing, then the possible ties will also grow, which will lead to fluctuation in density because the growth of total actual ties lags behind increase in membership. If participants work on increasing the number of people whom they speak to (degree), then density over time should continue to increase towards 1.

4.2.8 DIAMETER OF A LANGUAGE COMMUNITY

The diameter of the mitsqanaqan community graph in Figure 7 is the number of ties within the longest direct pathway between any 2 participants (Figure 7 is reposted below for simplicity). If the number of ties between participants is counted, the direct pathway between any 2 participants varies. Some pairs of participants have a direct pathway of 1 tie (e.g., 16 and 2); some have 2 ties between them (e.g., 11 and 3); but the longest direct pathway is 4 ties (e.g., 8 and 17). This means that any two participants are no more than 4 ties away from speaking to each other, and the diameter of this language community is 4.

The diameter of a social network does not measure language vitality; instead, its value lies in identifying an important characteristic of social networks. Social networks are characterized as having very small diameters in relation to the membership of a community (Jackson, 2008). This means that even with very large populations, people are only a few ties from knowing one another. The party game Six Degrees of Separation is an example of this where people attempt to find a connection to a famous individual within 6 ties. This relatively small diameter of social networks means that most participants tend to be relatively well connected, making it easier to relay information between them (Jackson, 2008). This includes transmission of linguistic information, such as pronunciation, vocabulary, grammar, and syntax. However, shorter paths of language transmission are not the only influence on language behavior, as discussed below. It is also affected by influencers, frequency, and intensity of communication, especially within clusters.

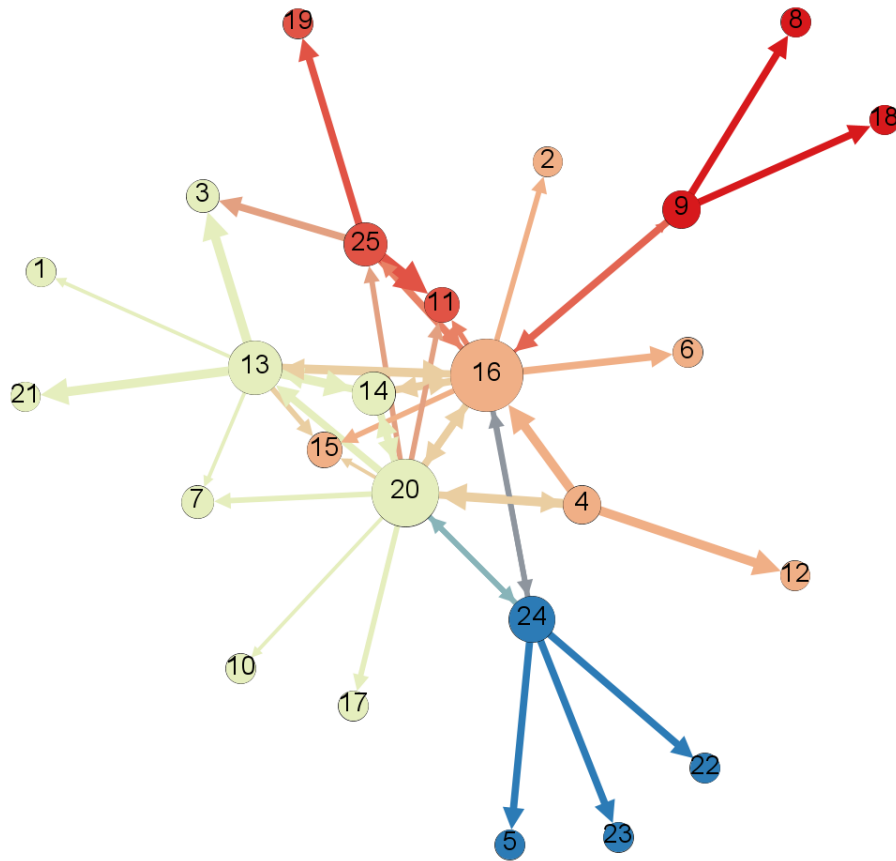


Figure 7: Network graph of the mitsqanaqan language community

4.2.9 OVERVIEW OF THE MITSQANAQAN LANGUAGE COMMUNITY

The statistics for the social network analysis of the mitsqanaqan language community is summarized in Table 13, below. Here, we can observe that the graph in Figure 7 has 25 participants and a total of 46 ties. The average degree means that on average, each person is communicating with 1.84 other people in mitsqanaqan. The density of a graph ranges from 0 to 1 where 0 means no one is speaking to each other, and 1 means that everyone is speaking to each other. The density of 0.077 in Table 13 represents a number that is less than 1-tenth (0.01) of the possible density. The diameter of the language community is 4, which is less than 1/6 of the membership of 25.

If over time, the statistics 1 through 5 increase and the diameter remains relatively small, then this is empirical evidence that the use of the mitsqanaqan language is expanding. If the statistics 1 through 5 decrease and the diameter increases, this is an indication that language socialization is decreasing,

meaning that people are communicating with fewer and fewer people in the language. Reasons for this such changes will vary from situation to situation. However, language programming should be structured towards getting people to speak to each other as much as possible. This means that primary instruction of classes should be focused on how to use language outside of the class, such as reclaiming domains, language nesting and conversation. This does not mean that language skills in pronunciation, grammar, syntax, and literacy should not be taught. However, it is essential that the content of these elements of speech only be taught as they relate to what the participant is using outside the instruction context and not be the primary focus of instruction.

Item	Statistic	Number
1	Participants (Network size)	25
2	Ties	46
3	Average degree	1.84
4	Average value	2.70
5	Graph density	0.077
6	Graph diameter	4

Table 13: Language Vitality Measurements for Mitsqanaqan

Beyond the statistics in Table 13, the network graph in Figure 7 also has visual features that inform us about the mitsqanaqan community. In Table 14 below, note that participant 16 has a degree of 17, followed by participant 20 with a degree of 15, and participant 13 with a degree of 10 (rows are in bold for clarity). These three participants are represented in Figure 7 as having the 3 largest circles. Noting these 3 participants have the largest degree does not reflect language vitality of the language community, but it does inform us as to who is well connected within the network. In social networks, individuals who have the most ties with others often have a greater influence. Such people of persuasion are termed 'influencers'. Within speech communities, they are key in affecting language socialization and linguistic comprehension and production. Limiting the number of influencers gives you greater control over language variation. However, doing so also hampers language expansion because the abilities of an elite few are limited by time and energy to speak to everyone, all the time, and about all topics. Therefore, it is aspirational to get as many participants speaking to others as possible, thereby increasing the community's average degree. If speaking is shared by all, then the language has more possibilities of being spoken by more participants for longer periods of time and in different contexts, thus increasing the health of the language. If mitsqanaqan language revitalization efforts continue to promote language socialization, then we can expect the degree of other participants graphed in Figure 7 to also increase.

Participant	In-degree	Out-degree	Degree
1	1	0	1
2	1	0	1
3	2	0	2
4	1	3	4
5	1	0	1
6	1	0	1
7	2	0	2
8	1	0	1
9	1	3	4
10	1	0	1
11	3	0	3
12	1	0	1
13	3	7	10
14	3	3	6
15	3	0	3
16	7	10	17
17	1	0	1
18	1	0	1
19	1	0	1
20	4	11	15
21	1	0	1
22	1	0	1
23	1	0	1
24	2	5	7
25	2	4	6
Total=	46	46	92

Table 14: Degrees: Mitsqanaqan Language Community

In addition, observe that the graph in Figure 7 has been generated with different colors. The coloring of the participants indicates where clusters of people are forming (Jackson, 2008). A 'cluster' is a group of participants that have dense internal connections compared to other external participants within the language community (Mishra et al., 2007). For example, the orange-colored participants – 15, 16, 12, 6, 4 and 2 – speak more to each other internally than externally, forming a social grouping within the language community. Participants of a cluster might include: close family members and

friends, coworkers, and members of a social group or club. Clusters tend to have persuasion over their internal members that can affect language use, variation, and behavior. Members of a tight-knit social group may motivate its members to speak with more intensity, create speech variation that evolves into a dialect, or inspire language use with domains not frequently used elsewhere in the language community. Knowing what social constructs form clusters helps us better understand what conditions language community vitality, language change, and other language behavior.

This brief overview is an example of how to use network analysis to assess a language community and its vitality. It does not cover all the many equations and algorithms network analysis has to offer but, instead, introduces basic statistics for language revitalizationists to use for language community vitality assessment. We are not suggesting that any one social network measurement or observation is better than another. Rather, we encourage vitality assessment use of all the social network techniques covered above, because each one informs us about a unique element of the community providing a more inclusive understanding of language behavior and health.

5. CONCLUSION

Our goal for this chapter was to introduce the idea that language community vitality is better analyzed through social network analysis. In addition to visual graphs, we introduced equations to assess characteristics of a language community, which we limited to those we felt most effectively and intuitively reflect language community vitality. Perhaps the most intuitive benefit of mapping any network is the visual analysis it suggests. The human eye is incredibly adept at discovering patterns, and most people will immediately see clusters and possibly patterns of relationship and flow of information between speakers. This strength, however, is also paired with weakness. Our urge to find patterns is so strong that we often find random patterns in networks that have no meaningful use (a phenomenon called 'apophenia'). This can lead to spurious analyses and incorrect conclusions. By reporting common measures that use common algorithms, we minimize the risk of a spurious visual analysis and unfounded conclusions, and we compare these common statistics across multiple language community networks.

This chapter does not intend to downplay the importance of the evaluation of language production and comprehension in terms of language skills and proficiencies. Assessing a language community as a social network first, though, allows linguistic fieldwork to narrow its investigation to those nodes within clusters that have a higher degree than those nodes that are in less dense areas. With more opportunities to speak to people, there is a potential that there is more speaking within clusters and that might correlate (or not) with an increase in language competencies and diverse functions.

REFERENCES

- Fishman, J. A. (1965). Who speaks what language to whom and when? *La Linguistique*, 1, 67–88.
- Fishman, J. A. (1991). *Reversing language shift*. Multilingual Matters Ltd.
- Gephi—The Open Graph Viz Platform*. (n.d.). Retrieved October 15, 2022, from <https://gephi.org/>
- Giles, H., Bourhis, R. Y., & Taylor, D. (1977). Towards a theory of language in ethnic group relations. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 307-348). Academic Press.
- Hinton, L. (2001). Language revitalization: An overview. In L. Hinton & K. Hale (Eds.), *The green book of language revitalization in practice* (pp. 3-18). Brill.
- Jackson, M. O. (2008). *Social and economic networks*. Princeton University Press.
- Krauss, M. E. (2007). Keynote—Mass language extinction and documentation: The race against time. In O. Miyaoka, O. Sakiyama, & M. E. Krauss (Eds.), *The vanishing languages of the pacific rim* (pp. 3-24). Oxford University Press.
- Labov, W. (1972). *Language in the inner city: Studies in the black English vernacular*. The University of Pennsylvania Press, Inc.
- Lee, N. H., & Van Way, J. (2016). Assessing levels of endangerment in the catalogue of endangered languages (ELCat) using the language endangerment index. *Language in Society*, 45, 271–292.
- Leonard, W. Y. (2011). Challenging ‘extinction’ through modern Miami language practices. *American Indian Culture and Research Journal*, 35(2), 135–160.
- Lewis, M., & Simons, G. (2010). Assessing endangerment: Expanding Fishman’s GIDS. *Revue Roumaine de Linguistique*, 55(2), 103-120.
- Light, R., & Moody, J. (Eds.). (2020). *The Oxford handbook of social networks*. Oxford University Press.
- Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics*, 21, 339–384.

Milroy, L. (1987). *Language and social networks* (2nd ed.). Basil Blackwell Ltd.

Mishra, N., Schreiber, R., Stanton, I., & Tarjan, R. E. (2007). Clustering social networks. In A. Botanto & F. R. K. Chung (Eds.), *Algorithms and models for the web-graph. WAW 2007*. (pp. 56-67). (Lecture Notes in Computer Science, Vol. 4863). Springer.

Newman, M. (2010). *Networks: An introduction*. Oxford University Press.

Rehg, K. L., & Campbell, L. (Eds.). (2018). *The Oxford handbook of endangered languages*. Oxford University Press.

Sapir, E. (1921). *Language: An introduction to the study of speech*. Harcourt, Brace.

UNESCO Ad Hoc Group on Endangered Languages. (2003). *Language vitality and endangerment*. UNESCO.

Ventureño language. (2022). In *Wikipedia*.

https://en.wikipedia.org/w/index.php?title=Venture%C3%B1o_language&oldid=1105908869

Zahir, Z. (2018). Language nesting in the home. In L. Hinton, L. Huss, & G. Roche (Eds.), *The Routledge handbook of language revitalization* (pp. 156–165). Routledge.