A SOCIAL NETWORK ANALYSIS OF IRISH LANGUAGE USE IN SOCIAL MEDIA

JOHN CAULFIELD

School of Welsh
Cardiff University

2013

This thesis is submitted to the School of Welsh, Cardiff University
in partial fulfillment of the requirements for the degree of PhD.

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (candidate) Date ……..

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed …… (candidate) Date ……..

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …… (candidate) Date ……..

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (candidate) Date ……..

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed …… (candidate) Date …….

ABSTRACT

A Social Network Analysis of Irish Language Use in Social Media

Statistics show that the world wide web is dominated by a few widely spoken languages. However, in quieter corners of the web, clusters of minority language speakers can be found interacting and sharing content. This study is the first to compare three such clusters of Irish language social media users. Social network analysis of the most active public sites of interaction through Irish – the Irish language blogosphere, the Irish language Twittersphere and a popular Irish language Facebook group – reveals unique networks of individuals communicating through Irish in unique and innovative ways. Firstly, it describes the members and their activity, and the size and structure of the networks they share. Then through focused discourse analysis of the core prolific users in each network it describes how the language has been adapted to computer-mediated communication.

This study found that the largest networks of Irish speakers comprised between 150-300 regular participants each. Most members were adults, male, and lived in towns and cities outside of the language's traditional heartland. Moreover, each group shared one common trait: though scattered geographically, through regular online interaction between core members they behave like communities. They were found to have shared histories, norms and customs, and self-awareness that their groups were unique. Furthermore, core users had adapted the language in new and innovative ways through their online discourse.

This study is the first comprehensive audit of who is using the Irish language socially on the web, where they are forming networks online, and how they are adapting the language to online discourse. It makes a unique contribution in re-imagining what constitutes an Irish language community in the context of the Network Society. In the process, it contributes to the growing body of sociolinguistic research into globalisation and local identity on the web.

Tá ceantar ag taisteal ón spéir
Tá comharsanacht suite ar mo mhéar

***

A locality is forming in the ether
A neighbourhood perches on my finger

Fiabhras by Seán Ó Ríordáin (1964)
Translated by Greg Delanty

ACKNOWLEDGEMENTS

I would like to acknowledge the support and generosity I received while carrying out this research. Without it this study simply would not have been completed.

Firstly, thanks to my supervisor Prof. Diarmait Mac Giolla Chríost for his support, advice and encouragement throughout the entire process. Thanks also to my co-supervisor Prof. Colin H. Williams, who was on hand for valuable advice at key moments in the study.

This research would not have been possible without the financial support from *an Roinn Gnóthaí Pobail, Tuaithe agus Gaeltachta* / the Department of Community, Rural and Gaeltacht Affairs, Ireland. I trust that the unique contribution this study makes to the sociolinguistic study of Irish was worth the investment.

I would like to acknowledge in particular the support given to me by the administrative staff of *Ysgol y Gymraeg* / the School of Welsh, Cardiff University. When called upon, they stepped in to help fund my training and fieldwork, and I am very grateful for their continued commitment to my research.

Finally, I would like to note the generosity of Prof. Kevin Scannell, Saint Louis University, and Dr Ruth Page, University of Leicester, who made their data and data capture software available to me. Without their shared expertise, the scope of this study would have been greatly limited.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I INTRODUCTION

A. THE NETWORK SOCIETY

1. Castells' 'Network Society'

Increasingly our lives can be understood in terms of our place in the invisible networks we share with others around us. This is what Manuel Castells calls the network society. According to Castells, our world constitutes a "space of flows" connected by a circuit of electronic exchanges, through the devices, telecommunications, computer-processing, broadcasting systems and high-speed transport links that increasingly organise our lives. In this space "localities become disembodied from their cultural, historical, geographic meaning and reintegrated into functional networks" (Castells, 1996: 375). He uses this metaphor of functional networks to characterise various aspects of modern life: from the shape of our cities to the flow of financial capital around the world. The advent of the internet[1] in particular has liberated communication flows across a global network of computer users.

The concept of the network society expands on the notion of 'community without propinquity' established by Webber in the 1960s as he sought to describe the changing social landscape of urban America. Webber (1963) recognised that as Americans became more socially mobile, and as communication became faster and cheaper, they were becoming more closely tied to communities of shared interest rather than traditional "place communities". To Webber, propinquity was no longer an accurate indicator of functional relations and, thus, "mere locational pattern [was] no longer an adequate symbol of order" (Webber, 1963: 49).

In 1986, Meyrowitz described the effects of electronic media as disassociating physical place and social 'place'. Though his study draws examples primarily from the telegram, telephone and television, and was written well before the internet gained popularity, his thinking still bears weight today. Before the arrival of electronic media, each place-situation was separated spatially and temporally from the next. In order to experience a culture or event, you had to be in a specific place at the right time. Place defined distinct situations because its physical boundaries "limited

[1] This study will refer to the internet, web, and email in their lower case, unhyphenated forms. This follows the Guardian style guide (http://www.guardian.co.uk/styleguide). The Guardian has a progressive view of capitalisation, changing from upper case to lower case spellings when the words are deemed to have entered common parlance.

perception and interaction" (Meyrowitz, 1986: 116). Travelling between places was a social 'rite of passage' as one moved physically and socially from one situation to the next. Today, however, electronic media have made it possible to feel present in social situations across the globe. Rolling news, instant messaging, 'live' feeds, all give us the feeling we are present in a multitude of situations. We can build relationships with others via the web without regard to any physical or social position.

Wellman (2002) charts this shift in modern society from traditional, homogenous, spatially-bounded local communities (what he terms "little boxes") to a society characterised by networked individualism. According to Wellman, the growth of computer technology and the availability of mobile communication devices have facilitated a transition from place-to-place to person-to-person connectivity, linking people with shared interests wherever they are. If the dominant theme of sociological study today is one of mobility and networked individuals, this has major implications for the way we understand society. As society transitions from traditional, local and geographically bounded contexts to a post-modern, globalised and interconnected world we are challenged as sociologists to devise new ways of understanding society in primarily social rather than spatial terms.

2. The Internet & the World Wide Web

One of the most significant developments in the shift to the network society has been the advent of the internet. In recent decades, the growth in popularity of the internet has connected us like never before, via a global network of interconnected computers, through what we generally call computer-mediated communication[2] (CMC). Moreover, the world wide web[3] brings together a network of billions of users publishing content and interacting with other web users at the click of a button. The web has liberated communication flows and created the potential for autonomous media production of unlimited diversity. Unlike traditional media, which were produced and distributed on a hierarchical one-to-many basis, media on the web is

[2] This term is not uncontested. David Crystal, for example, prefers the term electronically-mediated communication, which he uses to better include communication made electronically through devices outside the computer, e.g. mobile phones.

[3] The world wide web (web or www for short) is a system of interlinked hypertext documents. It is home to some of the most popular online activities: websites, blogs, social networking sites, video-sharing websites, etc. It is accessed via the internet: the global network of computers and other electronic devices interconnected via modems. The internet is also used in communication outside of the web; for example, for many email and instant messaging services, and for voice-over-internet services such as Skype.

characterised by more democratic production and reception. From its origins the internet was founded on the belief that many contribute to many, in a horizontal and egalitarian mode of operation (Kirshenblatt-Gimblett, 1996). Within this worldwide web of information we are given more power to choose how we express ourselves and to whom we listen. Admittedly some of the most popular websites and online resources are controlled by large media organisations, like Google, Yahoo! or Microsoft. But even on these sites web users may access and generate content free of charge and with limited editorial restriction.

The information we share via the web is "self-generated in content, self-directed in emission and self-selected in reception" (Castells, 2007: 248). However, among all this individual choice participants choose to cluster together in groups of similar interests, facilitated through the sociability of email, chatrooms, discussion forums, etc. While online, they interact with one another forming relationships that they may visit again and again over time. Their interaction through posting messages, linking to other sites and 'friending' other users creates an often explicitly linked social network, ideally suited for analysis.

3. Social Media

The emergence of social media in particular in the past decade has popularised the Internet as a site of regular and casual interaction between both acquaintances and strangers. Kaplan and Haenlein define social media as a new wave of internet-based applications that "allow the creation and exchange of User Generated Content" (Kaplan & Haenlein, 2010: 61). Whether it is the over 800 million active users on Facebook[4] or the Twitter users who publish over 340 million messages every day,[5] it is clear from the rapid uptake of these social network services that people feel increasingly comfortable engaging with friends, family and strangers via the web on a casual everyday basis. The formation of common interest groups through social media "has become easier now than at any other time in the history of communication" (Marsen, 2006: 176). Over time there emerge clusters of individuals using one or more of these websites to maintain regular contact with each other in what are casually termed 'online communities'.

[4] Data from LA Times, sourced 22/03/2012 http://latimesblogs.latimes.com/technology/2011/09/facebook-f8-media-features.html
[5] Data from Twitter blog, sourced 22/03/2012 from: http://blog.twitter.com/2012/03/twitter-turns-six.html

Three of the most popular social media worldwide are blogs, the micro-blogging site Twitter and the social networking site Facebook. Since these are the focus of this study, they merit further definition.

a. Blogs: A blog (short for web log) is a website on which individuals or groups of individuals publish posts on topics of interest to them. These posts are usually displayed in reverse-chronological order with the most recent post at the top of the screen, building up an archive of material over time. Blogging emerged as a popular form of online publishing in the late 1990s, growing rapidly in popularity during the 2000s. This growth was aided by the emergence of blog hosting services that provided platforms and templates for non-technical users to create and maintain blogs. Although much hype in recent years has been attracted to the rapid rise in popularity of other forms of social media, blogging still remains a powerful medium for sharing content and opinion with a global audience quickly and cheaply.

According to Blood (2002), there are three main types of blogs. *Personal journals* are blogs made up of posts where the author shares information about their own day-to-day lives and musings. This commonly takes the form of an online diary. *Filter blogs* reference external events and material, through hypertext links, videos, audio files, etc., engaging readers in content elsewhere on the web. *K-blogs*, or Knowledge blogs are those that build up content around specific topics or projects over time. Topics may vary widely depending on the interests of the individual bloggers.

According to Marlow (2004) there are four main types of social ties, or ways of connecting to others, in the blogosphere. These form a "social currency" linking bloggers and their audience in meaningful ways and making bloggers aware of who is reading and commenting on their posts (Marlow, 2004: 3). In different ways, these links form "conversations" (Herring *et al*, 2005: title) across the blogosphere between participants.

Blogrolls: A blogroll is a list of links to other blogs of interest to the blog author, usually displayed in the sidebar of the blog homepage. They are a common means of linking and navigating between blogs that share a theme or topic. A link within a blogroll indicates a "general social awareness" on behalf of the author (Marlow, 2004: 3), and acts as a form of endorsement, promoting other blogs of potential interest to their audience.

Permalinks: A permalink is a hypertext link embedded within a post, linking to material elsewhere on the web. They are used to form a "sort of distributed conversation" between blogs (Marlow, 2004: 3) as bloggers use permalinks to refer and respond to content posted elsewhere in the blogosphere. They represent more than just an awareness of the blogs being referred to, but active engagement in their content, however temporary that engagement might be.

Trackbacks: Trackbacks are automatic messages that are produced on a blog when that blog is referenced in the post of another. Both blogs need to be enabled with trackback functionality for this message to be generated. Trackbacks allow authors and readers to see who is discussing their content outside the comments on their own blog.

Comments: By far the most frequent form of interaction in the blogosphere, and arguably the strongest, is through comments (Schuster, 2004). These are replies to specific posts within the blogs that are contributed by readers. They represent a simple and direct way for bloggers and readers to interact with each other. Comments appear chronologically, and may build up threads of conversations on the one theme, as various readers and bloggers respond to each other's comments.

b. Twitter: Twitter is an online microblogging service in which users post messages up to 140 characters in length. Launched in the United States in 2006, the service grew rapidly in popularity worldwide: by its sixth year it had 140 million active users posting 340 million messages[6] (or 'tweets') every day. Unlike blogs, which can be set up independently and linked across different blogging services, Twitter users must create their own accounts on Twitter and interact according to the specific conventions of the service. They choose whose content to read by 'following' them (similar to 'friending' in Facebook). Twitter users have different levels of access to tweets on other accounts depending on whether they follow, and are followed by, them. Twitter is a popular social media globally and, as of 2012, its interface has been translated into 30 of the world's most spoken languages. Not surprisingly, given its origins in the USA, English is by far the most commonly used language on the site. Herring's (2009) linguistic analysis of over 36,000 messages posted to Twitter's public timeline during four hour-long samples over one day, showed that English was

[6] Data from Twitter blog, sourced 22/03/2012 from: http://blog.twitter.com/2012/03/twitter-turns-six.html

the dominant language, followed by Japanese and Spanish. In all samples, at least one in three tweets was written in English.

There are four basic forms of tweet. Firstly, users may publish *updates* to their Twitter accounts: short messages posted to the users' own Twitter feeds. These messages also appear automatically on the feeds of other Twitter users who have been accepted as followers of those accounts, enabling those followers to keep up-to-date on the activity of their favourite users. The number of followers varies widely from each account, with some celebrities attracting millions of followers worldwide.

The second means of communicating is through *@replies*. By posting the Twitter username of another individual at the beginning of a message and prefixed by the @ sign, that message is sent directly to that user regardless of him/her being a follower. If they are a follower it will show up on their Twitter feed and can be viewed in turn by all of their other followers. If they aren't following the @reply's author, it appears separately in a list of *mentions* only, and is not viewed by other followers. If the @username is included within the body of a tweet (i.e. is not the first word in the tweet) then that tweet is treated simultaneously as an update and a mention, that is, it is viewed by the author's followers and also sent to the person to whom it is addressed. Over time @replies between individuals may develop into a semantic thread, or conversation, albeit displayed reverse chronologically in the users' feeds and scattered among the individuals' other tweets.

Thirdly, Twitter users may choose to *retweet* content from other users. These messages usually begin with the letters RT (for 'retweet') followed by the username of the original message's author prefixed by the @ sign. This convention enables users to share, endorse or promote messages from elsewhere in the Twittersphere whilst acknowledging the original authors. Retweets appear in the feeds of the retweeters' followers, as well as the feeds of the messages' original authors.

Finally, Twitter users can send messages privately to other users in the form of *direct messages*. This is achieved by beginning the tweet with the letters DM (or just D) followed by the targeted account's @username. The recipient must be a follower of the direct message's author to receive it, and it does not appear in the feeds of the sender's or receiver's followers. Thus, direct messages remain private and outside the public discourse of the Twittersphere.[7]

[7] For a more detailed description of the various ways of communicating in Twitter, including examples of each type of tweet, see the *Mom This is How Twitter Works* website at http://www.momthisishowtwitterworks.

A convention common to Twitter and used variously across the four types of messages is the use of *hashtags*. These are words or phrases that are included in messages and prefixed by the # symbol usually to denote a theme or topic to which the messages refer. Page (2012) identified different communication strategies employed by celebrity, corporate and 'ordinary' accounts in the use of hashtags. While celebrity and corporate accounts used hashtags commonly to promote their visibility in the Twittersphere and to engage their audience in company names, slogans and products, 'ordinary' users were engaged in a more participatory culture of interaction around popular events, themes or topics. This latter use has been adopted by television and radio stations, for example, in engaging their audiences in discussing their programmes via Twitter using unique programme-related hashtags. Hashtags enable users to search more easily for specific content across the Twittersphere and to semantically group messages from users all around the world according to a specific search term. From a linguistic point of view, hashtags enable users to search more easily for tweets in specific languages.

c. Facebook: Launched in 2004, Facebook has grown rapidly to become the world's most popular social networking site.[8] Each Facebook user must register with the site, creating an individual account with a homepage where they can share text, images, audio and video with other registered users. Individual Facebook pages are usually private, requiring the owner to accept you as a 'friend' before you can access their content. New content posted to Facebook is displayed reverse-chronologically enabling users to keep track of the lives of their Facebook friends through the messages, photos and videos they share. Over time this builds an archive of content that can be viewed back to the beginning of the account.

There are many different ways to interact on Facebook. Facebook users can upload messages or content to their own or their friends' *walls*. The wall is a space on each Facebook user's personal homepage that is publicly viewable to all of their friends and where content uploaded by the user is displayed. Users can post messages directly to their friends' walls or comment on messages already posted there. Friends can also 'tag' content with the names of their friends (a way of associating messages,

[8] Data from *The Telegraph*, published August 2011 on: http://www.telegraph.co.uk/technology/google/8718580/The-ten-most-popular-social-networking-websites.html

photographs, videos, etc. with that user) to appear on the wall. Wall posts are stamped with the time and date of when they were published, and displayed in reverse-chronological order.

Friends can also choose to message each other privately, maintaining dialogue through Facebook out of sight of other users in a function similar to email. These private messages can have files attached. All private messages are displayed chronologically by username, building up threads of conversations over time between specific users. Multiple users can be addressed simultaneously with each response being made viewable to the group of addressees.

There is also a chat feature on Facebook through which friends can interact if they are both simultaneously logged onto their accounts. Users can see which of their friends are free to chat, and the short messages exchanged between them are kept hidden from other users. These chats are also archived and added to the other private messages between those users.

One simple way of interacting around content is by 'liking' it. All content uploaded to the wall has a 'like' link associated with it, which friends can click to show their appreciation. The number of people who liked the content is displayed, and their names can be revealed.

Facebook users can set up groups to invite others to join and interact around an event or theme. Closed groups are restricted to those who are invited by the group administrator, and their content can only be viewed by invited members. Open groups, on the other hand, can be searched for and joined by any other registered Facebook users, regardless of whether they are friends with the administrator or not. The messages posted to open groups can be read by both members and the public at large. However, to contribute to the interaction taking place there one must be accepted as a member.

This study focuses on the groups of individuals using blogs, Twitter and Facebook to interact through the Irish language. In analysing the discourse taking place there, it will refer to the various features of these websites and how they have been adapted to the Irish language – just one of the hundreds of languages with a presence online. To put this study in context, we now turn to the wider discussion around language on the internet.

B. LANGUAGE & THE INTERNET

1. The Multilingual Web

Evaluating linguistic diversity online is problematic given the scale of web output and the difficulty in identifying individual languages. Of the approximately 7,000 languages worldwide,[9] data on web presence exists for only the most widely spoken. A UNESCO report into linguistic diversity on the web estimated that 75% of web pages in 1998 were in English (Pimienta *et al*, 2009: 33). The dominance of English reflected the internet's birth and early popularity in the USA, as well as the language's initial 'default' status for international communication online (Warschauer, 2002: 62). By 2005, however, the percentage of web pages in English had fallen to 45% as citizens from non-anglophone countries had come online in increasing numbers (Pimienta *et al*, 2009: 33). This process continues today, aided by the worldwide growth in access to computers and internet technology, and increased access to web publishing in writing systems other than the Roman alphabet in a post-ASCII[10] internet. Figures from market research company Internet World Stats indicate that the web is still dominated by a small number of major languages, however. They estimated that in 2011 26.8% of web users spoke English, 24.2% spoke Chinese and 7.8% spoke Spanish, although it is unclear how these figures were deduced and how reliable they are. There is a marked difference between figures for the number of web pages in a given language, and the number of web users who speak that language. The Internet World Stats figures do not include data on multilingual speakers, for example. Despite these reservations, it is clear from the estimates of web pages and web users that the web is an increasingly multilingual platform, albeit dominated by a handful of major world languages.

The central dynamic affecting linguistic pluralism online, as described by Warschauer (2002: 71), is the contradiction between global networks and local identities. On the one hand, the internet promotes the use of English and other major world languages as *lingua francas* for global networking online. Danet & Herring

[9] Data from Ethnologue: Lewis, M. P., Simons, G. F. & Fennig, C. D. (2013), *Ethnologue: Languages of the World, Seventeenth edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com. Retrieved July 10th, 2012.

[10] ASCII, the American Standard Code for Information Interchange, is a character-encoding scheme, widely used to display text in the early days of the internet, that supports a character set based on the English language. Today, however it has been surpassed in popularity by schemes based on the Unicode system that supports diacritics, symbols and alphabets from a much wider collection of languages.

(2007) describe a linguistic hierarchy at play, with local languages often ceding to English and other dominant regional languages when individuals with different linguistic backgrounds interact online (see, for example, Swiss medical students' language choices on email lists in Durham, 2003). On the other hand, the internet provides a platform for interaction and publishing at a local level, on a one-to-one or many-to-many basis. This communication will frequently take place in a variety of languages and dialects depending on the language habits and competencies of the individual users and their intended audience.

One of the leading researchers in internet linguistics today is Jannis Androutopoulos, and the theme of local identity in a globalised web is at the heart of his study into language use online. For Androutsopoulos, globalisation is not a uni-directional process, whereby global discourses are adopted uncritically across cultures. Rather, his research explores the localisation and recontextualisation of online discourse to create new vernaculars on the web. Androutsopoulos bases his research primarily in non-English language settings on the web and explores within these spaces the dichotomy between the 'old vernaculars' of locally bound traditional ways of speaking and the 'new vernaculars' of digital culture (Androutsopoulos, 2011). In one study, for example, Androutopoulos analyses the 'spectacle' of two YouTube videos that have been dubbed into Bavarian German (Androutsopoulos, 2010). The multi-layered and complex language use exhibited in these videos points to a comfort on the part of the videomakers in playing with and mixing language style, variety and dialect on a global web platform. In another study Kytölä and Androutsopoulos (2012) carry out an ethnographic study of multilingual discourse in an online Finnish football forum. Despite most participants being Finnish speakers, they discovered that English in particular was found to act as a shared code. The researchers identified examples of code switching to English in complex lexical terms, formulaic phrases, and Finnish/English compounds. The results point to a fluid use of language among the forum users, with both medium and situation factors playing a role in informing language choice. The same curiosity that inspires Androutsopoulos to explore language use and change on the worldwide web, lies at the heart of this study and its focus on how the Irish language is evolved by its users in their interactions online.

2. Measuring language presence online

Estimating levels of language use on the web is notoriously difficult, and figures vary widely from study to study (Gerrand, 2007). It is not technically feasible to measure the total internet activity for a specific language directly, so in order to form an estimate for comparison, researchers, cultural organisations and private companies turn variously to three basic measures: *user profile*, *web presence* and/or *user activity*. Each approach measures different indicators of language use. Peter Gerrand (2007) explores the pros and cons of each method and the irregularities and inaccuracies in their results.

User profile is a measure of potential online language use based on the estimated number of active internet users in each language. It is most commonly deduced from the population of language users as recorded in national censuses and internet penetration levels in the countries where the languages are spoken. This technique is used by specialists like Global Reach[11] and Internet World Stats, with results quoted in UNESCO and OECD reports. One of the drawbacks of this approach, however, is that population statistics are often asynchronous and inconsistent, based on censuses from different years (and decades) and establishing language use through different means. Also, by measuring internet penetration at a national level, it fails to take into account varying rates of internet availability between regions depending on their level of economic development. This may overestimate or underestimate the use of minority languages associated with particular regions. Moreover, assumptions about online language use cannot be made based on crude census figures. The fact that an internet user speaks a language day-to-day does not mean that they will necessarily write in that language, for example (Gerrand, 2007). In the case of minority languages, internet users may struggle to find web resources in their vernacular language or an audience with which to communicate. As a result, they may turn to a more commonly used *lingua franca* when communicating online. Moreover, this approach struggles to account for multilingual users.

Web presence is a measure of the number or proportion of web pages written in each language across the web. It treats each web page in the sample equally, analysing the main language used. It is most commonly measured via search engines, searching for words or phrases across the chosen languages and comparing the

[11] Renamed Guava since Gerrand's article.

numbers of results. However, it may also involve analysing the languages used in randomly addressed websites sourced from across the web. The advantage of the web presence approach is that it estimates language use based on evidence from existing web pages. However, the number and sources of results will vary widely depending on which search engines and search terms are used. If on the other hand a web trawl is carried out, snowballing outwards through links between websites, then the chosen starting points may bias the sample and the languages used within them. In addition web searches will not take into account communication via the internet that remains hidden from the web, like e-mails, instant messaging, etc. Furthermore, without reference to how frequently the web page is accessed, the results can only measure a potential, rather than actual, use of the languages under study. Finally, the methods used to identify different languages may be flawed. Searches may only include languages written in roman alphabets, or may omit minority languages that are more difficult to identify. For example, a Babel study in 2007[12] was limited to only 15 of the most commonly used languages online. As Wright (2006: 191) cautions, reporting in any quantitative way on the state of the web is a "dangerous enterprise". Even if statistics on web presence can be reliably obtained, before any research is published the situation will have evolved considerably. The incredible growth of the internet, and the short life span of many sites, means that any quantitative data on web presence quickly becomes inaccurate. Wright concludes that an exhaustive, quantitative linguistic survey of websites would be "pointless".

This leaves the *user activity* approach, which measures the actual use of a language on the internet. As it is unfeasible to carry out a full-scale linguistic audit of the internet due to its massive scale, user activity is most often measured on a specific technology or genre, or within a well-defined online setting. Gerrand refers to three "micro studies" using this technique, analysing the languages used within an email list for Swiss medical students (Durham, 2003), a bulletin board system for students in a Catalonian university (Climent *et al*, 2003), and a European Union discussion forum (Wodak & Wright, 2006). Depending on the size of the sample, measuring user activity may vary from human coding of content, to computer-aided searches for key words, phrases or diacritical marks associated with a given language. The advantage

[12] Babel (1997), "Web languages hit parade – June 1997". Online: http://alis.isoc.org/palmares.en.html Retrieved January 7, 2010

of this approach is that it measures language as it is actually used, in the content individual users produce online. The disadvantage is the length of time it takes to analyse content within different users' messages. This inhibits the scale of the research. As in the web presence approach, if computer software is used to speed up the process by searching for specific linguistic markers, then inconsistencies may arise from the search terms used and the ability of software to recognise different alphabets and minority language markers. Despite these shortcomings, small-scale studies of language users' activity are the most common approach to analysing minority language use online, with researchers frequently turning to ethnographic techniques to understand the social processes taking place within groups of web users.

3. Minority languages online

Despite the dominance of English and other widely spoken languages online, the web represents an enormous opportunity for minority languages to be written, read and heard by a global audience that makes traditional media reach – by newspapers, radio and television – look miniscule by comparison. In his treatise on language death, David Crystal (2000: 141) listed six significant factors for language revitalisation, one of which states:

> "An endangered language will progress if its speakers can make use of electronic technology".

Of course not all language communities will have the resources and the determination to make the transition to the internet. UNESCO lists 2,470[13] endangered languages in its Atlas of the World's Languages in Danger (Moseley, 2010). These comprise languages and creoles from around the globe that share varying levels of endangerment ranging from 'vulnerable' to 'critically endangered', with all of them facing uncertain futures. Half have just a thousand speakers or fewer. For the majority of these languages, concentrated in remote parts of the developing world, the creation of an online presence may seem like an unnecessary and unattainable goal in the immediate battle for survival. For the most critically endangered languages, the internet may play a basic role in simply recording and documenting the last surviving speakers before the languages' extinction.

[13] It lists a further 254 extinct languages.

The internet affords less critically endangered languages spoken in regions with wider internet access a significant platform for language use. In his study into both the threats and opportunities afforded to minority languages through the internet, Cunliffe (2007) extols the web's potential to form active communities of minority language producers working collaboratively. Not only does the web offer a platform through which the indigenous production of minority language material can be made widely and cheaply, it also provides the opportunity for web users to engage in a global language community online, no longer tied to a geographic location. Moreover, there are potential opportunities for languages to be used in new and innovative ways online, liberating minority language use from its traditional contexts. This may be particularly attractive to younger users, offering new motivations for them to use their ancestral language via new technologies (Crystal, 2011: 86). Cunliffe (2007) cites Langer's (2001) suggestion that between 1999 and 2001 the number of Welsh language web pages per speaker increased by almost five times. However, he goes on to argue that without further analysis this figure gives little insight into the content of these pages or the frequency and way in which they are used. As Cunliffe (2007: 139) states, "the absence of detailed figures and analysis of online minority language presence makes it difficult to define what constitutes a successful online presence and what the stages are in achieving such a presence".

Despite the enthusiasm among sociolinguists for the potential of electronic technologies to promote the use of minority and lesser-used languages, they continue to be significantly under-represented in studies of user profile and web presence. Gerrand cites work by Guinovart (2003) and Mas (2003), which provide some web presence estimates for the Catalan, Galician, Basque, Welsh, Frisian and Faroese languages, using the multilingual search engine AllTheWeb. However, he questions their reliability, owing to the fallibility of search engines and the untestable nature of the results. Deere (2011) cites Greffenstette and Nioche's (2000) data on the web presence of 32 primarily European languages, including Irish. Using the Altavista search engine it estimated the number of words in each language online by combining the search results of certain words from each language with the frequency with which those words occurred in the language. The results pointed to the overwhelming dominance of English, which represented an estimated 78% of content found in the 32 languages. Irish, in comparison, comprised just 0.083% of content, and Welsh was the least well represented language with 0.013%. Deere (2011) discusses flaws associated

with the design of the study. However, the overall trend in the results from Greffenstette and Nioche, Guinovart, and Mas confirms the peripheral status of minority and regional languages online, under the dominance of English and, to a lesser extent, other world languages.

In comparing the activity of minority language use online one might turn to traditional methods of measuring ethnolinguistic vitality in the real world. Giles, Bourhis and Taylor describe ethnolinguistic vitality as "that which makes a group likely to behave as a distinctive and active collective entity in intergroup situations" (Giles, Bourhis & Taylor, 1977: 306). Vitality is imagined through a language community's perceived strengths and weaknesses in a number of dimensions, including demographic strength, social status and institutional support. According to Giles *et al*, the more vital the ethnolinguistic group was perceived to be in these dimensions the more likely it was to survive as a distinct linguistic collectivity. Conversely, weaker groups were expected to assimilate linguistically or cease to exist as distinctive groups.

The challenge for the online sociolinguist is to translate these real world factors into an online context in which the size and structure of online communities may be difficult to gauge, their social status within a world wide web of billions of users may be impossible to assess, the institutional support they enjoy may be difficult to determine, and the users' individual language competencies may be difficult to measure in a text-based computer-mediated context.

In their methodology for assessing language vitality and endangerment,[14] UNESCO states clearly that a language's vitality cannot be measured by one factor alone but rather as a combination of different factors. Used primarily to gauge levels of threat to the world's lesser-used languages, UNESCO's nine measures include traditional demographic factors such as the absolute number of speakers, their proportion to the total population, and the age of speakers (i.e. the rate of intergenerational transmission); attitudes towards the language on an individual as well as governmental and institutional level; and the language's use in different media and genres: responses to new domains and media, availability of materials for language education and policy, and types and quality of documentation.

[14] *Language Vitality and Endangerment* (2003) sourced from the UNESCO website on 24/07/2012 from: http://unesdoc.unesco.org/images/0018/001836/183699E.pdf

Degree of Endangerment	Grade	New Domains and Media Accepted by the Endangered Language
Dynamic	5	The language is used in all new domains
Robust/active	4	The language is used in most new domains
Receptive	3	The language is used in many new domains
Coping	2	The language is used in some new domains
Minimal	1	The language is used only in a few new domains
Inactive	0	The language is not used in any new domains

Table 1. UNESCO grades of response to new domains and media.

One of the criteria used by UNESCO categorises languages according to six different grades depending on their response to new domains and media. They caution that language communities that fail to embrace the "challenges of modernity" risk becoming irrelevant and stigmatised (Pimienta *et al*, 2003: 11). Table 1 shows these grades, ranging from dynamic to inactive, according to how actively the language communities have adapted to new domains, such as schools and new work environments, and media, such as radio and television. These grades are likely to differ greatly between minority languages in developed countries and those in remote regions of poor and underdeveloped countries. As the UNESCO report cautions, even when languages have been adapted to new media there may be significant differences in the number of newspapers, radio stations or TV channels, for example, or the amount of air-time and content available in each language. This study will refer to the grades in Table 1 in comparing the online vitality of Irish and other European minority language.

Although this study refers to a number of different minority languages, the Irish language – and how it is used interactively online – is the primary focus for analysis. As such, and to put later analysis in context, the history of the language and its contemporary minority status is described below.

C. THE IRISH LANGUAGE

1. The Irish language

Irish – called Gaeilge in the standardised form of the language, and often anglicised to Gaelic or Irish Gaelic (not to be confused with its sister language Scots

Gaelic) – is an indigenous language of Ireland, and a member of the Celtic family of languages. The Irish language's contemporary status as a minority language[15] belies a rich history of literature and song and a cultural heritage of continued language use reaching back to the fourth century AD when a Proto version of the language was first spoken on the island (Mac Giolla Chríost, 2005: 64). Indeed, Irish was the language of the majority of people living on the island of Ireland for centuries. Like Welsh and Scots Gaelic, however, the Irish language has suffered a history of decline and peripheralisation under the dominance of the English language. The decline of Irish occurred for a variety of socioeconomic reasons, among them: oppression of the language by an English-speaking political and legal system; the decline and eventual collapse of the Gaelic chieftain society culminating in the so-called 'Flight of the Earls' in the early 17th century; emigration and rural depopulation from the language's traditional heartland, particularly heightened during the Great Famine in the 1840s; and a growing association of English as the language of modernity and socioeconomic advancement (Mac Giolla Chríost, 2005). With the creation of the Irish Free State in 1921, successive governments have pledged to preserve and promote the language. The 1937 constitution established Irish as the Republic of Ireland's "first official language", and in 2007 it was made an official language of the European Union.

Irish is a compulsory subject in state schools in the Republic of Ireland at both primary and secondary level. This helps account for the relatively large number of people – 1.77 million in the 2011 census[16] or just over 40% of the population aged three or over – who claim they 'can speak Irish'. When asked about the frequency of language use the census data indicates that over 600,000 people (14% of the population aged three or over) speak the language daily, but that the majority of daily speakers are children who use the language exclusively within the education system. This leaves approximately 94,000 people – children and adults – who use the language every day outside of the education system, comprising just over 2% of the population aged three or over. Of these, the most visible clusters of Irish speakers live

[15] The term 'minority language' refers here to the number of speakers of Irish. It is the preferred term for this study for the way it describes those languages that "are dominated politically and economically by numerically larger communities within a particular state" (Cormack, 2007: 2). Other commentators (see Ó Laoire, 2005) argue that because of its official status within the Irish State it cannot be considered as a 'minority' language. Kelly-Holmes (2011: 44) avoids this criticism by describing the language as a "privileged, minoritised language".
[16] Data available from http://www.cso.ie.

in the Gaeltacht: the name given to those regions in Ireland where Irish is still used as an everyday community language.

The census results paint an image of the Irish language that inspires both optimism and pessimism for its survival. It would appear that there are increasing numbers of 'speakers', but of these only a small fraction use the language on a daily basis. Indeed, the broader outlook for the language is a mix of both positives and negatives. Some commentators point to a perceived increase in positive attitudes towards the language in Ireland in recent decades (see Dawson, 2010). This can be evidenced, for example, in the remarkable growth of Irish medium education across the island. However, legislation and government support have failed to reverse the decline of the language in the communities where it has traditionally been spoken, and the preservation of the linguistic heritage of these 'Gaeltacht' communites is under severe threat. A recent study of language use among young people in these regions concluded that without a major change to language-use patterns, Irish would be unlikely to remain the predominant community and family language in even the strongest Gaeltacht regions for more than fifteen to twenty years (Ó Giollagáin *et al*, 2007: 27).

2. The 'Gaeltacht'

As a concept, the Gaeltacht deserves some further definition here because it represents an established way of talking about communities of Irish speakers. The term Gaeltacht originally meant an Irish-speaking people in general (Ó Laoire, 2005). However, since 1926 it has had a very specific geographic and demographic definition. In that year, *Fíor-Ghaeltacht* – 'true' Irish-speaking community – status was granted to those electoral divisions in the Irish Free State where at least 80% of the community were Irish speakers (Ó Giollagáin *et al*, 2007). This new definition came from a macro language policy to achieve "some kind of regional delimitation" of Irish speakers for administrative purposes (Ó Laoire, 2005: 293). The result, according to Johnson (1997: 159), was "to fix the Gaeltacht in space and […] to define socio-linguistic policy in static geographical categories". In 1956, the Department of the Gaeltacht was established to promote the cultural, social and economic welfare of the Gaeltacht. At this point the boundaries of the Gaeltacht were redrawn to reflect the new statutory definition of the Gaeltacht as:

> "specified areas, being substantially Irish-speaking areas and areas contiguous thereto which, in the opinion of the Government, ought to be included in the Gaeltacht with a view to preserving and extending the use of Irish as a vernacular language" (Ministers and Secretaries (Amendment) Act, 1956).[17]

This removed any demographic prerequisite for Gaeltacht designation, with a more ambiguous requirement for the regions to be 'substantially Irish-speaking'. Since 1956, the Gaeltacht has comprised rural pockets in seven of the Republic's 26 counties: Donegal, Mayo, Galway, Kerry, Cork, Waterford and Meath. Although the census shows that the number of Irish speakers living in the Gaeltacht is higher now than at any point in the past fifty years, the percentage of Irish speakers living within its boundaries has continued to decline, and the language's status as a community language remains under threat. Just over 100,000 people live in the Gaeltacht today, with roughly 66,000 aged three or over identified as Irish speakers. According to data from the most recent census on the frequency of language use, approximately 36% of Gaeltacht dwellers aged three or over use the language on a daily basis outside the education system. This is an average across all of the Gaeltacht regions, however, and in reality the language is more vibrant in some communities than others.

Although initially used by the State as a language planning tool, the Gaeltacht has come to define the Irish speaking community. Moreover, it engages people in a way of talking about a sustainable, viable body of Irish speakers. The Gaeltacht is seen to "occupy a particularly important place in the historical development of Irish society's conception of its identity, real and imagined" (Ó Riagáin, 1992: 102). Its primary geographic setting, in rural pockets on the western fringes of the island, promotes a romanticised idyll of remote local community life, a synecdoche of Irish identity. Its visitors are drawn by an:

> "idealisation of rural life, of 'traditional' life-styles in the Gaeltacht, an extraordinary static vision of Gaeltacht society, timelessly in tune with the elemental values of the Irish people; the repository of the linguistic elixir of Irish nationhood…

[17] Available online on 'Acts of the Oireachtas' website: http://www.acts.ie/en.act.1956.0021.1.html. Accessed 14/04/2013.

> from which the rest of the country could draw sustenance" (Ó Tuathaigh, 1990: 11).

Hundreds of schoolchildren are *sent* to the Gaeltacht each summer to improve their Irish and to experience Gaelic culture first hand. In a practical, administrative sense the Gaeltacht has been the focus of state planning and investment in preserving the Irish language. It sets aside a place where the continuity of the language is seen to be passed safely from generation to generation. This is its statutory role. From a legal standpoint, it is used by activists to campaign for language rights issues (Ó Laoire, 2005). But, in a symbolic sense it has awarded a special status to speakers who are reared and live inside its boundaries.

3. Irish outside the Gaeltacht

As Cronin (2005: 12) argues, our strong association between language and geography perpetuates:

> "the erroneous notion that Irish is uniquely a language of concrete placedness rather than free-floating abstraction, more soundscape than think-tank".

This is exemplified by the historical shift in the definition of a Gaeltacht from an Irish-speaking *people* to an Irish-speaking *place* (see Ó Torna, 2005). Moreover, it acts to tie the fate of the language to the fate of the Gaeltacht (Cronin, 2005), and its very existence excludes many Irish speakers living outside the Gaeltacht from the fruits of state investment and from the recognition it awards them. Its crude geographical definition of language use overlooks the use of Irish outside the Gaeltacht and underestimates "the complex spatial interconnections" between the Gaeltacht and other places (Johnson, 1997: 188). This association between the Irish language and the geography of the Gaeltacht contrasts markedly with Castells' description of the contemporary network society. It begs the question, how are Irish language communities to be reimagined in a society where localities are, in Castells' words, becoming increasingly "disembodied from their cultural, historical, geographic meaning" and reintegrated into functional networks?

As Ó Laoire (2005: 293) states "naturally occurring variations in social patterns of language use do not allow for neat demarcation of speech community" and there is evidence of significant networks of Irish speakers living in towns and cities across the island. Ó Broin (2010) suggests that the growth of urban Irish-language radio stations

is evidence that Ireland's towns and cities are "reaching a critical mass of second-language Irish speakers who want their own media". Moreover, there has been a significant growth, over the past 30 years in particular, in Irish-medium education at nursery, primary and post-primary levels outside the Gaeltacht. This not only acts to produce a new generation of more competent Irish speakers but also creates focal points for Irish language use in the community outside the Gaeltacht context. Mac Giolla Chríost, for example, cites Ó Riagáin and Ó Gliasáin (1979) in their study of Irish in Dublin in the 1970s and their assertion of "the centrality of Irish-medium schools to the workings of Irish-speaking networks in the city" (Mac Giolla Chríost, 2005: 203).

Non-Gaeltacht Irish speakers differ from their Gaeltacht counterparts, in that their networks have never been "sufficiently numerous to form a readily identifiable and easily visible speech community" (Ó Laoire, 2005: 277). It is argued that the capacity for these networks to reproduce themselves "is severely restricted by their size, thin distribution, and transitory character" (Ó Laoire, 2005: 285). Outside the Gaeltacht, Irish represents a private language of family and close acquaintances rather than a community language at large. In the words of Michael Cronin (2005: 9), *"is treibh dofheicthe iad":* they are an invisible tribe. However, if these speakers were somehow more visibly networked, they could represent an exciting new type of Irish language community. Indeed, the Gaeltacht Act 2012 passed by the current Irish government calls for the designation of 'Irish language networks' in communities outside the Gaeltacht. These would be communities that 'supported' the Irish language and that agreed to implement Irish language plans for their areas. It is unsure what the advantages of this new designation would be, but presumably communities recognised in this way would benefit from state support in promoting the Irish language locally. As the introduction to the Act on the Department of Arts, Heritage and the Gaeltacht website explains, it is envisaged that through this new shift in focus under the Gaeltacht Act 2012 "the Gaeltacht will in future be based on linguistic criteria instead of on geographic areas which has been the position to date".[18]

In fact some communities and groups outside the Gaeltacht's boundaries have already adopted and adapted the term 'Gaeltacht' to their own ends in creative ways, in the context of also adopting some of the linguistic legitimacy and authenticity

[18] Retrieved on 14/01/2013 from the Department of Arts, Heritage and the Gaeltacht website: http://www.ahg.gov.ie/en/20YearStrategyfortheIrishLanguage/GaeltachtAct2012/.

associated with the term. Mac Giolla Chríost (2012) describes the creation of 'Gaeltacht huts' in the early 1970s in the Long Kesh prison in Northern Ireland where Republican prisoners taught the language to each other. The emergent community of prisoners who learned, conversed and wrote in the Irish language while within the prison system would come to be known as the 'Jailtacht' – a portmanteau of the words 'jail' and 'Gaeltacht'. There are two other examples of the term 'Gaeltacht' being adopted by communities in Belfast city: the Shaw's Road Gaeltacht, a cluster of Irish-speaking homes in the south-west of the city; and the city's Gaeltacht Quarter, which represents an unspecific area centred on the Falls Road in the west of the city where a number of Irish language services and organisations are based. Since 2007 a 'North American Gaeltacht' has been designated in Tamworth, Ontario, Canada, which acts as a centre for Irish language enthusiasts in North America. None of the examples above fulfil the geographic and demographic requirements for Gaeltacht designation in the Republic of Ireland. Rather, they point to a more fluid interpretation of the term. These 'neo-Gaeltachts' challenge us to reimagine Irish-speaking communities in a totally new context, as networks of people and services using Irish explicitly as their habitual language of communication, in locations that are primarily English-speaking and far-removed from the traditional rural Gaeltacht setting. If the Gaeltacht defines those regions where the Irish language has been passed from generation to generation in an unbroken link through time, then the neo-Gaeltacht describes networks of Irish language users who have chosen to revive the language as an everyday community language in places where this link has been broken or, in the case of the North American Gaeltacht, never existed.

4. Irish and the Media

Traditional media has played a key role in bringing Irish language content to a national audience. Since the 1930s, various national Irish language newspapers have been published on-and-off under different titles (Ní Chartúir, 2002). This tradition, beginning with *An tÉireannach* in 1934, continues today with *Foinse* being published as a weekly supplement in the English language *Irish Independent* newspaper. Since it was founded in 1960 the public service broadcaster *Raidió Teilifís Éireann* (RTÉ) has included some Irish language content in its radio and television broadcasts. Most significantly, as a result of agitation from Gaeltacht community groups, the foundation of *Raidió na Gaeltachta* in 1972, under the aegis of RTÉ, created for the

first time an all-Irish language station that could be heard across the Republic of Ireland. With one studio in Dublin and others located in each Gaeltacht county, the station acts to "bring the far-scattered Gaeltachtaí within earshot of each other and break down provincial rivalries and prejudicial attitudes to regional dialects" (Robinson, 2011: 279).

> "The radio station has had the effect of reducing the importance of geographical separation of the Gaeltachts by connecting areas cut off from each other spatially, and thus creating a network where Irish speakers can exchange news and views that extend beyond the strictly local" (Johnson, 1997: 193).

Another all-Irish radio station, *Raidió na Life*, broadcasts to Dublin listeners alone, while *Raidió Fáilte* broadcasts in Irish to a Belfast audience. In 1996, a national Irish language television station *TnaG* (*Teilifís na Gaeilge*), later rebranded *TG4*, was founded in the Connemara Gaeltacht of west Galway. Beginning with just five hours of service a day (Ní Chartúir, 2002), output has expanded to a 24-hour service. From its beginning, TG4 constructed its mission as "that of being an entertaining television station rather than in terms of the discourse of endangerment" (Kelly-Holmes, 2011: 45), and it has a much less rigid language policy in its output compared to *Raidió na Gaeltachta*. TG4 switches between home-produced and imported content, in a mix of languages, and the policy of indigenising global formats, subtitling all recorded content with English language subtitles, and broadcasting Irish language promos, weather and continuity announcements before and after English language programmes has brought Irish language content to new audiences outside the Gaeltacht.

Today, the internet is an important platform for Irish language media and content. *Foinse* publishes material regularly online that can be accessed across the world. Similarly *Raidió na Gaeltachta, Raidió Life, Raidió Fáilte* and *TG4* broadcast live online to a global audience, and provide archived podcasts and broadcasts of their programmes on their websites. These are supplemented by other Irish language websites, created by media organisations, educational institutions, commercial businesses, community groups, and individuals, bringing a wide range of Irish language content to web users across a variety of online genres in what is colloquially termed 'Gaeltacht 2.0' (Lenihan, 2011: 50). Despite this, however, it remains the

norm for websites across Ireland to be published primarily and, most often, monolingually in English. There are no legislative provisions for the language on the web (Lenihan, 2011), and language policy development in this area is limited to a few small-scale studies of Irish schoolchildren (see Foras na Gaeilge, 2009; Fleming & Debski, 2007). In some cases users themselves have created localised versions of popular web services, such as *Facebook* and *Wikipedia* (see Lenihan, 2011: Deere 2011) in Irish, through a process of crowd-sourcing. While Irish language content on these sites may form a tiny fraction of the information being shared across them globally, they remain a valuable resource for Irish language web users, and provide a platform for the language to be used in totally new contexts.

D. THE AIMS OF THIS RESEARCH PROJECT

In her foreword to *A New View of the Irish Language*, Nic Pháidín challenges us to look to the future of the Irish language using a compass that points in multiple directions both real and virtual, "from geographical communities to cyber-based networks, from the Aran Islands classroom to the google-user of *focal.ie* inside the Arctic Circle" (Nic Pháidín, 2008: ix). This study answers that challenge in part by exploring how and where networks of Irish language users coalesce online in what might be called online communities. It seeks to understand how these networks are constructed within the non-geography of cyberspace, and to analyse how the Irish language itself is adapted by users, in the new context of on-screen computer-mediated communication.

To these ends, the focus now shifts to how other researchers have conceived and analysed online communities in their work. This will comprise (a) a review of literature on the concept of the 'online community' itself, including its advocates and its critics, (b) an introduction to social network analysis as a tool to understanding social structure, including examples from leading sociolinguistic study into online social networks, and (c) a review of research into computer-mediated communication, and an introduction to Susan Herring's approach to computer-mediated discourse analysis.

II DEFINING COMMUNITIES ONLINE

This section introduces the reader to the concept of online community. In the process, it describes how other researchers have attempted to define online communities and to describe both the social processes and discourse taking place within them. No universally accepted definition of online community exists, and the term is taken to mean different things by different people (Preece & Maloney-Krichmar, 2003). It is therefore worth exploring how scholars variously define the term 'community' in ways that stand up to academic rigour. As such, the literature reviewed below has a methodological slant – one that will inform how this study approaches identifying Irish language online communities in an empirical manner and describing the discourse features that set them apart.

A. ONLINE COMMUNITIES

1. The Online Community Debate

The term 'community' is a problematic one, loaded with values and with widely subjective connotations. The computer-mediated communication (CMC) lexicon is littered with phrases that infer close relationships between users: 'friends' meet in 'chatrooms', sign 'guest books', share 'private messages', etc. Yet, how can we be sure that these new forms of affiliation are more than just casual encounters? Two opposing schools of thought frame the debate about community online. To some, technologies like the internet represent a threat to traditional, 'real life' community interaction. Spending time with these technologies distracts us from being more active, face-to-face, in our physical localities. To the critics, communication between members is too casual and shallow, their topics of interest too narrowly defined to be classed as proper communities. Given the fluid membership, reduced social accountability and lack of shared geographical space that categorise online groups in general, there is some scepticism that community can exist in a virtual context (Herring, 2004: 344); moreover, it is argued that automatically labelling all online groups as 'communities' risks overextending the term "to the point of becoming meaningless" (ibid.). Writers such as Stratton (1997), Lockard (1997) and Snyder (1996) challenge the "myth" of online communality:

> "To accept only communication in place of a community's manifold functions is to sell our common faith in community vastly short" (Lockard, 1997: 225).

According to Lockard, we project feelings of community onto our online communications in a desperate attempt to feel connected in a world where traditional forms of community are disintegrating around us. This process is promoted by internet companies, like AOL, who benefit from consumers subscribing to their various online offerings in the search for communality. Stratton expands on the theme by arguing that the web empowers Americans (the majority of web users at the time of his writing) to feel connected to an artificial community much the same as watching a soap opera does. Both lure their audiences with nostalgia for the small town America of the past. Meanwhile, Snyder derides the ease at which web users can unsubscribe or drop out of their virtual communities at the click of a button. This is not true of "real communities" where people are forced to engage with each other because of their shared geographical proximity (Snyder, 1996)

To dismiss the existence of online communities is to suggest that communities in the 'real world' are somehow more tangible, more straightforward. In fact we use the term community casually to describe a whole range of collectivities from the local to the international. There is a vast difference in scale between the local community hall, the farming community, the gay community, and the European Community (Thurlow *et al*, 2004: 97); and yet the fact that they are commonly labelled communities implies that their members share some sort of implicit bond, a reassuring sense of togetherness. The term 'community', therefore, has a symbolic use invoked by people as they learn to be social and to interpret symbols that compose a shared culture (Cherny, 1999). As Benedict Anderson states:

> "All communities larger than primordial villages of face-to-face contact (and perhaps even these) are imagined" (Anderson, 1983: 15).

Anderson was specifically addressing the history of nationalism, but his thinking goes to the heart of what defines a community. He argued that a nation is a 'community' because it is conceived and promoted as such, and its citizens encouraged to believe that they share a "deep, horizontal comradeship" (ibid.: 7); it is 'imagined' because even in the smallest of nations no citizen will ever meet or know all of his fellow citizens. Nonetheless in their minds' eye lives "the image of their communion" (ibid.:

15). Communities are distinguished from each other not by their genuineness – no community is more genuine than the next – but by "the style in which they are imagined" (ibid.). The common faith that we share in 'real world' communities, so valued by Lockard, might therefore be imagined in other settings where individuals feel a bond to others outside the confines of what traditionally constitutes a community.

The question remains: is an online network a community? Castells (2010) answers yes and no. The confusion lies in how community is defined. Yes, they are communities, but not in a traditional sense. They work in a different plane of reality with different patterns of communication and interaction.

> "They are interpersonal social networks, most of them based on weak ties, highly diversified and specialized, still able to generate reciprocity and support by the dynamics of sustained interaction" (Castells, 2010: 389).

The dichotomy between being connected to a collective group online while simultaneously disembodied from your physical surroundings is a problematic one (Wilson, 2007); one that requires new ways of conceptualising ties between people. This challenge is summed up well by Sandy Stone in her description of electronic spaces as "incontrovertibly social spaces in which people still meet face-to-face but under new conditions of both 'meet' and 'face'" (Stone, 1991: 85). It is argued by early CMC scholars, like Rheingold, and Baym, among others, that communication technologies help unite people across the globe like never before, creating the opportunity for *new*, exciting communities of people connected by shared interests rather than geographical proximity. They favour a definition of community based on sociability rather than locality; one that doesn't rely on comparisons to traditional communities of the past.

> "CMC liberates interpersonal relations from the confines of physical locality and thus creates opportunities for new, but genuine… communities" (Rheingold, 1993: 5).

Moreover, these technologies can enhance and reinvigorate existing offline communities as they migrate online, or conversely initiate face-to-face contact among people in the 'real world', thus becoming "woven into the fabric of offline life rather than set in opposition to it" (Baym, 1998: 63). As Rheingold states:

> "in traditional kinds of communities, we are accustomed to meeting people, then getting to know them; in virtual communities, you can get to know people and then choose to meet them. In some cases, you can get to know people whom you might never meet on the physical plane" (Rheingold, 1987: 4).

The external contexts within which they are partially embedded must be considered separately for each community.

Writing in the mid-1990s about the relationships between users of an MUD,[19] Sherry Turkle explored the complex identities that form through CMC:

> "In the real-time communities of cyberspace, we are dwellers on the threshold between real and virtual, unsure of our footing, inventing ourselves as we go along" (Turkle, 1995: 10).

She shared Rheingold's hope and optimism for the potential for CMC to connect people in new and innovative ways. Fifteen years later, and after an explosion in the use of online social media, Turkle's optimism has waned. In her book, *Alone Together* (2010a), she laments the ways in which new technologies have isolated us from real human interactions. She describes how spending too much time online can at best distract, and at worst disconnect us, from the world around us. In her own words, "we have invented inspiring and enhancing technologies, yet we have allowed them to diminish us" (Turkle, 2010b: 17). Turkle's warnings are worth heeding. Too much time spent alone and online can be extremely unsocial and isolating. Moreover, one need only look at the aggressive practices of flaming[20] or the disruptive and often hostile practices of trolling,[21] common in online discussion forums, to see how anti-social some online interaction can become.

It is clear that not everyone interacting online is doing so as part of an online 'community'. Many studies have, however, shown groups of web users interacting online in socially meaningful ways. The variety of ways in which they describe online communality indicates that there is no simple definition. Online communities have

[19] An MUD (Multi-User Domain, also Multi-User Dimension, Multi-User Dungeon): a type of multiplayer real-time online game, which takes place in a virtual world described primarily in text. MUDs were an early popular form of online community and were the focus of much sociological study in the late 1990s.

[20] Flaming describes the, sometimes deliberate, use of hostile and insulting rhetoric in interactions online, frequently including profanity.

[21] Trolling involves the disingenuous posting of inflammatory or off-topic content deliberately intended to provoke an emotional reaction from other web users.

been shown in studies sharing affiliation around a wide spectrum of passions and needs: groups of fans who share music of very specific tastes (Lysloff, 2003), people passionate about politics and current affairs (Hine, 2000), diasporic communities connecting with their homeland (Mitra, 2000; Miller and Slater, 2000), communities who share languages or dialects (Siebenhaar, 2006), professionals sharing advice and services through technology (Uncapher, 2007), people who want to share their problems around illness or relationships (Bakardjieva, 2007). The challenge for researchers is how to identify online communities in culturally meaningful terms, while simultaneously grounding their analysis in empirically observable phenomena (Herring, 2004).

2. Towards a Workable Definition

In her study of community dynamics in the virtual world of an MUD, Cherny (1999) brings together a bewildering list of definitions of community. Communities are characterised to varying degrees in different studies by: norms of behaviour among members and group-wide goals (Arendt, 1958); a shared sense of tradition and commitment to one another (Bellah *et al*, 1985); a shared common interest through which they interact over time (Bilson, 1995); a shared knowledge and learning through participation (Lave & Wenger, 1991); a gossip circuit (Elias, 1974); a common good among members with punishment for bad behaviour (Smith, 1992); a spatially contingent structure within whose boundaries members exercise control over social processes (Jonassen, 1959); a shared understanding of their community's boundaries among members (Cohen, 1985); a social differentiation within society (Suttles, 1972); and a common imagery and ideology, including icons, symbols and slogans (Hummon, 1990). So how are online communities to be imagined? And how might we evaluate their existence on the web in an empirical manner?

The most cited definition of communities in an online context comes from Howard Rheingold. He describes them as:

> "social aggregations that emerge from the Net when enough people carry on […] public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace" (Rheingold, 1993: 5).

Rheingold's definition is simply an expression of the types of relationships we expect to observe in an online community. It is significant in creating a benchmark by which

we might characterise communality across different groups and genres on the web. It is nevertheless a subjective definition, open to interpretation and not empirically testable. A researcher might justifiably ask: how long must a discussion forum exist before it is deemed 'long enough' to be a community? How does one blog exhibit 'sufficient human feeling' more than another? How many people using a chatroom is 'enough people' to comprise a legitimate online community? Depending on the nature of their site of interaction – blog, social network site, chatroom, discussion forum, etc. – users exhibit different levels of anonymity, exclusivity, synchronicity and social cues that define human feeling differently (Joinson, 2003).

In an effort to make identifying online communities more operationalisable, Herring (2004) distils the literature on features of virtual communities into six distinct criteria. They represent various aspects of the "sociability, support and identity" posited by Wellman (2001) as being at the heart of online communities. Not all online communities will exhibit all of these features, and some of the criteria may be more useful than others as potential indicators of virtual community in different circumstances. In distinguishing these criteria, Herring aims to make the process of labelling online groups as 'communities' a more empirical one, based on objective assessment. She describes them loosely under six headings and without much detail, preferring to cite various studies in which they have been identified differently. They are outlined below with additional comments in some cases from Baym's (2010) study into personal connections online. Baym focused on the qualities that distinguish online communities from the users' points of view: how they are imagined by the members themselves, and how a sense of community is made to resonate with members in tangible ways. These qualities overlap with and compliment Herring's criteria. The criteria are as follows:

- Participation: Online communities should comprise a core of regular participants, actively engaged with each other through what Herring calls "self-sustaining participation" (Herring, 2004: 364). Participation can be measured over time, identifying core participants in the community as those with the most frequent posts and the highest rate of responses received to their messages.
- Shared history and culture: Over time, and through regular interaction, an online community develops its own sense of purpose and culture, with a

variety of norms of behaviour. A shared history can be identified by analysing an online group's archived messages. Herring draws on a number of previous linguistic studies into various online communities to indicate how shared culture is made explicit in the context of on-screen and text-based CMC. These include the use of group-specific abbreviations, jargon and language routines, as well as the users' choice of language, register and dialect. She points to netiquette statements, existence of FAQs, and reactions to inappropriate conduct as examples of shared values and norms. Baym also describes how communities may be distinguished from each other through their habitual practices, positing shared innovative language use as a key identifier in this regard. In identifying how shared cultures are formed within the group dynamic, she distinguishes between those groups where norms are explicitly and hierarchically imposed in moderated online communities, and those unmoderated groups where they remain implicit and negotiated.

- Solidarity, support and reciprocity: Reciprocity is perhaps the easiest of these phenomena to measure, through an analysis of interactions between community members for turn initiation and levels of response. Solidarity and support between members is rather more nuanced, and Herring points to the use of verbal humour and acts of positive politeness as examples of evidence of these traits in other studies into online communities. As Baym argues, the support shared within communities may be distinguished by the types of resources their members exchange. The social capital shared between members acts to bond them together in the exchange of advice, information, feedback and empathy; and in turn gives community members the feeling of being needed. This shared emotional support pertains to the 'human feeling' Rheingold identified in his definition of online communities.
- Criticism and conflict: These are rather more unexpected characteristics of online communality but are a natural part of group sociability. Herring identifies conflict as speech acts between community members that can be seen to violate positive politeness. These violations suggest the breaking of community norms and values. She highlights methods of resolving conflict between individuals as an important characteristic of online communities.

- Self-awareness: This describes awareness among members that their group is a distinct entity, and different from other groups. This is made explicit in the way members refer to the community. Herring describes the use of 'us vs. them' language that can be witnessed, for example, in how community members express the way things are done in their group, as opposed to in other groups. According to Baym, members of online groups often imagine them as shared places. Though most online participants do not share a geographical locale, the online communities in which they interact are often imagined and described in ways that constitute "a metaphorical sense of shared space" (Baym, 2010: 77). This is evident in the use of spatial expressions in talking about websites. Thus, MySpace becomes a "social entertainment destination". Fan pages become "the home of" Twenty20 cricket (http://www.cricket20.com), or "your destination for" the English Premier League (http://www.goal.com). 'Visitors' grow accustomed to the sites' individual design as they "enter" and "leave" a website (Bergs, 2000: web page). They come to know who and what to expect within these spaces, and relationships are understood within the specific boundaries of that community.
- Roles and hierarchy: Group members may inhabit distinct roles within the group, developed over time. A discussion of hierarchy within online communities may be based around perceptions of leaders and non-leaders within the group dynamic. Herring suggests that these can be deduced from participation patterns within the group and by analysing speech acts by group members. Baym cites Welser *et al* (2007) in distinguishing here between 'answer people', those that respond to messages frequently but never initiate them, and 'discussion people', who both initiate and respond to messages. Groups that exhibit relatively advanced stages of community may share distinct governance structures and forms of group rituals, although Herring argues that these should not be taken as part of any basic definition of online communities.

The stronger online networks are characterised in these ways, the more they satisfy Rheingold's definition of online communities. The rather subjective terms used in his definition correspond to the more clearly outlined qualities and features in Herring's list of criteria. If we look again at Rheingold's definition there are a number of distinct

clauses that must be satisfied. Online communities are "*social aggregations* that emerge from the Net when *enough people* carry on […] *public discussions long enough,* with *sufficient human feeling,* to form *webs of personal relationships* in cyberspace" (Rheingold, 1993: 5). Firstly, the emergence of "social aggregations" with "enough people" may be assessed through participation patterns, identifying how many individuals are involved and how regularly they communicate with each other. The frequency with which group members interact may vary greatly between the groups' core and peripheral participants. Secondly, "long public discussions" suggests a shared history within the group that can be analysed, as Herring suggests, by searching back through the groups' archives of messages. The ease with which this can be achieved depends on the persistence of messages on the website in question, as well as the levels of access and privacy settings within the specific online group. Evidence that groups have a shared past is also seen in the self-awareness members have of their group being a distinct entity. Moreover, the emergence of specific roles among participants is further evidence that the community has developed over time. Thirdly, Rheingold's "sufficient human feeling" pertains to a number of qualities mentioned by Herring and Baym. These include evidence of shared solidarity, support and reciprocity, and the emergence of a shared culture among group members with expected norms and values. Criticism and conflict between individuals and, in particular, means of conflict resolution indicate a more than just casual relationship between group members. Finally, "webs of personal relationships" can be explored through analysing participation patterns and the reciprocal exchange of social capital between members. Personal relationships will not form equally across the entire network of participants. Instead, one would expect to find different rates of interaction and stronger ties formed between core participants compared to those on the periphery of the community.

The studies cited in Herring's examples use a range of methodological tools, each tailored to the specific qualities being explored and the individual groups under study. The challenge for CMC scholars is to devise ways of analysing user relationships in various online settings so that sense of community might be measured and compared. The two approaches chosen by this researcher in exploring the groups of Irish language users online are social network analysis and computer-mediated discourse analysis. Specifically, social network analysis (SNA) focused on the structural

qualities that define online communities. Computer-mediated discourse analysis (CMDA) on the other hand focused on the social processes taking place within the groups as evidenced in the exchanges between group members on screen.

B. SOCIAL NETWORK ANALYSIS

1. Social Networks

Social network theory makes the basic assumption that we do not exist in isolation from each other or as free members of individual groups, but that we share social networks in which every element is "somehow elementarily linked to other elements in the system" (Bergs, 2006: web article). These social networks consist of webs of social exchanges that people share between each other, be they family, friends, neighbours, work colleagues, acquaintances or strangers. Each social network consists of individual actors as nodes (or points, or hubs), and relations between actors as ties (or links, or edges, or lines) between nodes. These relations can take many forms; they may represent the exchange of information, resources, money or support, for example (Garton *et al*, 1997). As an individual's multiple relations with others are observed and as, in turn, these individuals' relations with others are observed, so the social network stretches and grows outwards.

Not all nodes in a network are equal; rather, the density, multiplexity and strength of interpersonal ties between each node affect the network's structure and their position within it.

Density: Put simply, network density is a measure of the extent to which all nodes are connected to each other. If the people you know and interact with also know and interact with each other, you are said to share a dense (or closed) network (Wardaugh, 2002). If however, they do not share ties in common, it is said to be a loose (or open) network. Social network analysts use the term 'degree' to describe the number of nodes with which an individual node has a direct tie. The higher the degrees of nodes in a network the denser it will be (Scott, 2000). The degree of each node will vary widely, depending on the size and inclusiveness of a network. Thus, network density is a relative measure, based on the rate of interconnectedness between nodes within the boundaries of a particular network. Furthermore, the way density is measured will depend on what types of ties are under study. One study might measure the density of

all nodes connected in *any* way, while another might take the direction of the ties into account. In the latter case each node's indegree, i.e. the total number of other nodes that have ties directed towards it, and outdegree, i.e. the total number of other nodes to which it directs ties, will be analysed separately (Scott, 2000) indicating levels of reciprocity within the network.

Multiplexity: A network's multiplexity increases when members of a social network share more than one tie with one another. For example, if a group of friends lives close by and works with each other as well as socialising together, they share a multiplex social network (Britain & Matsumoto, 2005). Conversely, someone who is only tied in a network through one type of social connection is said to have a uniplex network link. Multiplex ties may be shared through various media across a variety of offline and online contexts. Multiplexity affects how ties are measured in different studies, with some weighting ties differently according to how multiplex they are. Dense and multiplex ties within a small group of nodes are characteristic of a close-knit social network structure (Milroy & Milroy, 1992). This will appear as a cluster of nodes in social network visualisations.

Tie strength: Within their network of friends, a person might distinguish between their best friends or close friends, with whom they meet regularly, and what they might call distant friends or acquaintances, with whom they share less frequent and less emotionally engaged contact. Just as individual nodes in a social network will have different numbers of connections, these connections may also be of varying strength: what we might call variations in the intensity of ties (Milroy & Milroy, 1992). According to Granovetter (1973: 1361), the strength of a tie is "a combination of the amount of time, the emotional intensity, the intimacy […], and the reciprocal services which characterise the tie". Strong ties are characterised by more frequent contact; weak ties are characterised by "less frequent, more transient and more incidental" contact between people (Paolillo, 2001: 186, citing J. Milroy, 1992). Strength of tie may be assessed by a combination of factors, including frequency of contact, reciprocity of contact, duration of the relationship, number of media used for interaction, the intimacy of the tie and kinship between individuals (Haythornthwaite, 2002). These are relative measures, with the factors chosen and the methods through which they are measured varying across networks and studies. A person is likely to share more weak ties than strong ties, as weak ties require less effort to maintain than the time and energy invested in maintaining stronger ties. This is increasingly the case

in more mobile, urban societies (Milroy & Milroy, 1992). Haythornthwaite (2002) coined the term "latent ties" to describe all of the potential relationships that could exist within a network but have not yet been activated. Unacquainted friends of friends would fall into this category. They may remain latent ties for life, or may be activated to form a tie as circumstances change.

The strength of a tie changes over the duration of a relationship, growing stronger as people get to know each other better; and weakening, or ending, as their motivation for contact or their medium of contact disappears. Ties can be created and maintained through multiple media, and strong and multiplex ties might share contact through a mix of face-to-face encounters and telephone calls to text messages, letters, emails and online social networking. The type of tie that people share affects their ways and means of expression as well as the motivations and desires for communication. Haythornthwaite (2002) argues that people who share strong ties can influence each other to adapt and expand their relationship through other media, branching out into other forms of communication. They are less reliant on any single medium and, therefore, their network remains robust in the face of changes in media choice and access. In contrast, people with weak ties are less motivated to adapt their relationships to other media, and remain reliant on means of communication established by others. Thus, weakly tied networks may disintegrate and communication may become disrupted when changes occur to their medium of communication.

Taken together these three dynamics of density, multiplexity and tie strength have fundamental influences on how social networks act and are sustained. Milroy and Gordon (2003: 118) claim that if a personal network consists chiefly of strong ties that are also multiplex and within a relatively dense structure, then such a network has the capacity to "support its members in both practical and symbolic ways". As Boissevain (1979: 393) states, in analysing social networks we must ask "questions about whom is linked to whom, the nature of that linkage, and how the nature of the linkage affects behaviour". Many academics turn to social network analysis (SNA) to help understand this linkage.

2. Social Network Analysis

Social network analysis (SNA) is a discipline of social science that seeks to make sense of the patterns or regularities in relationships across social networks

(Wasserman & Faust, 1994). In other words, it is concerned with the nuts and bolts of social structure. Social network analysts use relational data to measure the many contacts, ties and connections between members of a social network (Scott, 2000). Its strength lies in the potential to describe social structure through patterns of interaction between people when other reliable cues to social structure are lacking (Paolillo, 2001). In contrast to macro-level constructs such as class, gender and ethnicity, SNA allows the researcher to examine "the specifics of local practice and local conditions" within local social categories and locally constructed ties (Milroy & Gordon, 2003: 116). It strives to determine the social structure within these networks based on empirically observable exchanges to determine how and what resources flow between nodes. It is hoped that observing these micro-level exchanges will in turn shed light on macro-level social phenomena.

a. Social Network Analysis and Sociolinguistics

SNA is used as a tool in many academic fields, including economics, organisational studies and information science – anywhere phenomena can best be described in terms of the patterns of relations between people. It is a particularly appropriate tool for the study of sociolinguistics, as language only exists in its exchange between people. Whether it is written, spoken or signed, the language we use is learned, shaped and changed by and through the connections we share with others. Sociolinguistic study based on SNA is concerned with how the density, multiplexity and strength of these connections affect language use within social networks.

Daming, Xiaomei and Wei (2008) outline the advantages of SNA in approaching sociolinguistic study. Firstly, SNA may help the researcher gain access to private networks that might otherwise be inaccessible. By using a 'friend-of-a-friend' approach to finding and accessing speakers, speech communities and speech events, the researcher may make further contacts within the social network. This may enable the researcher to observe a much larger number of people over a shorter period of time. Secondly, it offers the researcher a specific perspective that is embedded within existing human social relations. Once the social network is identified, the researcher knows the boundaries within which their research can base its analysis. Allowing the social network to dictate the range of the research can help avoid claims of bias. Any

conclusions made from the research will be stated only in terms of the social network in question.

Milroy and Gordon (2003) outline the specific advantages of SNA to variationist sociolinguists. It enables the researcher to study small groups of speakers, where the differences in language use may not be discernible in terms of any greater social class index. SNA can be used to study variation at the level of individual speakers, through their interpersonal ties, avoiding the need to talk about individuals in terms of predetermined social groups. Furthermore, SNA can elucidate the social dynamics effecting language variation and change. This enables the researcher to explore language variation through social factors developed over time.

b. Social Network Analysis and the Web

SNA is particularly suited to the study of computer-mediated communication (CMC). In online networks, social background information may be scarce or unavailable, and SNA enables researchers to examine web users through the observable social structure of their online interactions (Paolillo, 2001). Clusters of web users may be described more practically in social network terms rather than using more subjective or discursive language. Conversely, CMC is also particularly suited to SNA. Relatively large amounts of data about web users' communication can be collected and analysed, compared to traditional research into face-to-face interactions. For some researchers the ability to tap this data is causing a new computational social science to emerge that "leverages the capacity to collect and analyse data with an unprecedented breadth and depth and scale" (Lazer *et al*, 2009: 721). As they navigate and communicate online, web users leave footprints behind in different ways. These footprints can be traced to identify the ties between different users in a tangible way. With publicly accessible online content, we can know exactly "who talked to whom about what for how long" (Bergs, 2006: web article). This is rarely, if ever, achievable with research into spoken language on such a large scale.

The most basic defining feature of CMC, its most fundamental structural property and that which marks it apart from all other forms of communication is the hypertext link (Crystal, 2009). Hypertext links are the means by which people navigate the web, 'clicking through' webpages as they go. This process allows them access to billions of webpages in a controlled and repeatable manner, making sense of the seemingly haphazard infrastructure of the web (Gibson *et al,* 1998). Hypertext links are used in

various creative and interactive ways to navigate, promote, explain and endorse aspects of online communication, all the time connecting the intricate network of resources that comprises the web. Moreover, the hypertextual structure of the web makes linking explicit (Herring *et al*, 2005). As they navigate through the web, users leave behind a trail of hypertext links in their wake. This information is of value to analysts who seek to understand online behaviour, including webmasters who set up hit counters to quantify and analyse visits to their sites, and internet marketers who gauge the success of online advertising by their rates of click-through. Hypertext links can be used to extract a meaningful structure for online communities that might otherwise lack any explicit logical organisation (Gibson *et al*, 1998). This is done through an analysis of the link topology to expose any core-periphery structure or hierarchy within the community. Furthermore, social network analysts use the hypertext links shared between web users as indicators of relationships and 'conversations' (Herring *et al*, 2005), and in describing those relationships, be they reciprocal or not.

Another important feature of CMC is the use of individual online identities in the form of usernames, user profiles and avatars. These are key to identifying individual users and their online activities for researchers conducting online social network analysis in participant-observer roles and without access to website hit counters. On many sites, users must create a username to access and contribute to online social groups like blogs, chatrooms, newsgroups or social networking sites. Once registered, these unique usernames allow users to 'log in' to their chosen websites. Commonly in social networking sites, for example, users must be invited to become an online 'friend' to access personal information. Without this status, users can only view very basic information. Importantly, usernames appear each time individuals communicate with others, thereby identifying the source and recipient of interactions online. The usernames may identify the person's actual name, may be a nickname that hints at their gender and background, or may be completely anonymous. Despite the optional anonymity they nevertheless help mark users apart from each other. In this way they can be used by social network analysts to structure their SNA around distinct individual nodes.

Where users are not required to log in to view content they may visit online social groups purely as voyeurs, solely *viewing* material left by and commented on by others, in an act known as 'lurking'. This may be the height of a person's

involvement, choosing never to upload material or interactions of their own. This may leave the author and other users unaware of the extent of their audience. Hit counters can be used by webmasters to identify the IP addresses of visitors to their sites. The resulting data indicates lurkers' rough geographic location and how frequently they visit the site. However, without engaging these visitors in identifying themselves by logging in under a username or posting their own material online, the lurkers, and their media habits, remain completely anonymous. Without any activity on which to base their connections within the network, lurkers remain excluded from any social network analysis of users.

Haythornthwaite (2002) makes three key assumptions about ties in online social networks. Firstly, she argues that the characteristics of ties hold in online networks as they do in offline networks. Thus, we can expect to see strong ties between people who share more frequent, intimate and multiplex interaction. Moreover, it is assumed that weak ties play an important role in diffusing innovation across online networks as they do offline. Secondly, it is assumed that online exchanges resemble offline exchanges in the ways in which they impact ties. Interaction mediated through online environments supports a range of emotions – from emotional support to flaming – just as it does in offline environments. Individuals' reactions to these emotions affect the strength of ties they share with others. Thirdly, Haythornthwaite argues that the number and type of exchanges between individuals is determined by the tie they share, and not by the environment in which they communicate. Individuals sharing strong ties are expected to branch out into different methods of communication, potentially both online and offline, as the desire to communicate in different ways increases. In contrast, individuals sharing weak ties are unlikely to feel propelled to adapt their interaction to other media. To gain the measure of a computer-mediated tie, therefore, it is important to analyse if and how it is maintained offline or in different web genres online.

c. Social Network Visualisation

SNA has seen a notable surge in interest since the early 1990s. Wasserman *et al* (2005: 1) account for this increased interest as a "realization in much of behavioural science that the 'social contexts' of actions matter". Another factor leading to its popularisation is the increased availability and power of computer software to analyse social network data and create visualisations. This software can extract and analyse

data from large samples, and create visualisations of the intricate webs of relations that connect social networks. Lines of varying lengths reflecting tie strength create clusters of different densities based on dyadic data. Nodes and connectors can be coloured and shaped differently to highlight different phenomena of interest to the researcher. Social network visualisation has greatly increased the potential for identifying patterns in social network structures that may otherwise remain hidden in data, and has made SNA more accessible to new audiences.

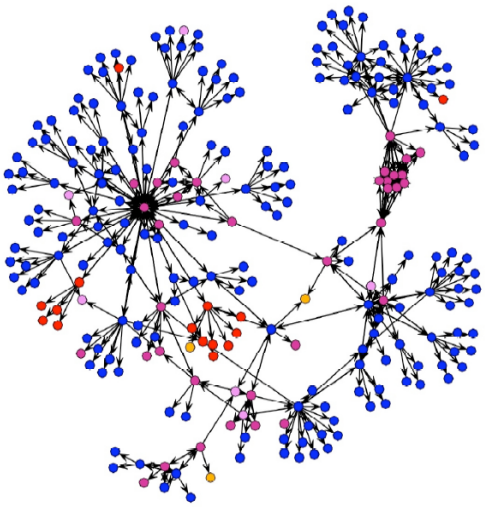


Fig. 1 Example of a social network visualisation, including nodes, connectors and clusters, from Herring *et al* (2007).

Social network visualisations may help reveal different social structures within networks (see Fig. 1): clusters of tightly-grouped nodes, poorly-connected nodes on the periphery, nodes acting as 'bridges' between different clusters (Herring *et al*, 2007), or indeed social isolates or 'islanders' whose networks are separated from others (Daming *et al*, 2008). Examining these structures helps us imagine networks of individuals in spatial, as well as social, terms.

Some sociolinguists use SNA techniques to help visualise and understand linguistic patterns in social networks. In his study into language use in an English-language Indian Internet Relay Chat (IRC) group, Paolillo (2001) analysed the presence of chosen linguistic variations among participants' messages. These included abbreviating the words 'are' and 'you' to 'r' and 'u' in a process known as rebus spelling; marked non-standard spellings in the replacement of the letter 's' with the letter 'z' at the end of plural words; use of obscenity; and code-switching from English to Hindi within their texts. He tested Milroy and Milroy's (1992) hypothesis

that vernacular variables are expected to be used more frequently by those members of a social network that share stronger ties, while more legitimised and standard norms should be evident among those with weaker ties. If the social processes at work in virtual community formation on IRC are like the social processes of geographically based communities, then we should expect to find patterns of variation similar to those observed by Milroy and Milroy (1992) in their study of Belfast communities. Paolillo constructed a social network by counting the frequency and direction of interactions between each pair of participants in the chatroom. Within this structure he analysed the presence of his chosen linguistic phenomena. The combined information identified cliques of chatroom users sharing linguistic variation. Two of these variations, the use of obscenity and Hindi code-switching, appeared to be shared more frequently between participants with stronger ties. However, the other variations were less strongly connected to tie strength. Use of 'r' and 'u' was widespread across the sample, leading Paolillo to claim that, contrary to his original hypothesis, they should be considered more as standard variables. The study also concluded that non-standard use of 'z' at the end of plurals was, unexpectedly, more commonly shared by participants with weaker ties. Paolillo accepted that the 'r', 'u' and 'z' variations must be explained by other sociolinguistic factors and admitted that his original hypothesis based on tie strength did not accurately predict the distribution of linguistic variables. He stated that in each case there was a need to take into account the "larger social context of the observed interactions in which the variables are used" (Paolillo, 2001: 209). Despite the inconclusive results, Paolillo's study shows the potential for SNA to at least form a structure for sociolinguistic research in well-defined communities where interactions can be quantitatively analysed. Once SNA identifies clusters of nodes within a social network, other methodologies might be employed to describe why these clusters exist and the social processes taking place within them.

3. Social Network Analysis in the Work of Susan Herring *et al.*

Susan Herring is a pioneer of the SNA of CMC. Her research analyses online communities 'from the bottom up', as she puts it in the title of one paper (Herring *et al*, 2005), using data and content analysis from web users to describe their patterns of relationships. She uses this analysis to answer various research questions about the social and linguistic processes occurring within online social networks. The following

two studies were conducted by Herring in conjunction with other academics from the School of Library and Information Science (SLIS), Indiana University Bloomington.

'Conversations in the Blogosphere'

In 2005, Herring *et al* carried out research to investigate the extent to which blogs were interconnected across the blogosphere (Herring *et al*, 2005). The blogosphere is characterised as a "massively distributed but completely connected conversation covering every imaginable topic of interest" (Marlow, 2004: 1). The researchers conducted a multiple ego-centric approach social network study of four randomly-selected blogs and analysed their links to other blogs through three degrees of separation. They analysed the resulting social network, both quantitatively and visually, to interpret its structure and to make judgments on its interconnectedness. Links were identified in each case by counting hypertext links to other blogs, conducting a search for mentions of other blog names or bloggers' personal names throughout the entire blog, and by reading comments to the 20 most recent entries to find references to other blogs. The study also differentiated between what they called 'A-list' and 'non A-list' blogs; the A-list being the most popular blogs as listed by three different blog aggregators in their Top 100 blogs. With this data in hand, they gauged the degree of overlap, distance between nodes, reciprocity of linkage and the number of inbound versus outbound links between each blog. Though involving only a tiny fraction of the millions of existing blogs, it was assumed that the results from this sample would be suggestive of some properties of the blogosphere as a whole.

In general, the social network that emerged from the study was quite interconnected. Highlighting some of the researchers' observations here will, I hope, demonstrate the potential for SNA to elicit interesting and surprising results about CMC. Firstly, Herring *et al* concluded that linking to other blogs on one's own blog was likely to lead to reciprocal linking. They found, however, that this was more likely to occur among non A-list blogs, who had fewer inbound links to start with compared to their A-list counterparts. Citing Gibson *et al* (1998), Herring *et al* drew parallels between the rate of inbound and outbound links and a blog's status within its group of adjacent nodes. They noted that blogs who have little or no outbound links but receive many inbound links may be considered 'authorities'; while blogs with a high degree of outbound links may be considered 'hubs'. They also found that linking between non A-list blogs increased the likelihood of those bloggers entering a

'conversation' with each other in their entries and comments. Herring *et al* identified three main clusters of varying sizes and densities among their sample. Two of these clusters – which they labelled Catholic blogs and Homeschooling blogs – reflected the topics of interest of two of the randomly chosen blogs at the start of the study. These blog clusters were not recognised by the original quantitative analysis, but later emerged through visualisation. The third cluster centred on A-list blogs of a primarily political nature. Both the Catholic and Homeschooling blogs were more densely clustered than the A-list blogs, suggesting that there existed close-knit groups of bloggers reciprocally linked by shared minority interests and remaining under the radar of popular blog aggregators. The study also made a distinction between filter blogs, which consist primarily of hypertext links to content elsewhere on the web; and personal journal blogs, which consist of the bloggers' own musings on life around them. Herring *et al* suggested that filter blogs play a more central role in the social network of the blogosphere. As filter blogs contain more links to other blogs they make more fruitful places to focus ego-centric approaches to SNA.

'Language Networks on LiveJournal'

In 2007, Herring *et al* carried out another SNA of the blogosphere, this time from the perspective of language use across a network. Again they took a multiple ego-centric approach, albeit from a much larger starting sample of 24 blogs divided equally between four languages of primary use: Russian, Japanese, Portuguese and Finnish (see Fig. 2). These blogs were all hosted by LiveJournal.com, a US-based blog hosting service, which at the time of study offered its interface in 32 different languages. This time, connections to other blogs were analysed two degrees outwards. Again, hypertext links and comments were used as evidence of connections between blogs, but in this study they were supplemented by a service on LiveJournal.com whereby bloggers display messages from their selected 'friends' and include information in their profile about 'friends-of-a-friend'. It was found that the hand-coding of blogs, where the researchers physically looked through each blog to identify the primary language in use, was the only reliable way to gather language data. The resulting data was used to visualise the social network with each node coloured differently according to the language of primary use. The data and visualisations were used to answer the following research questions: how robust are the networks formed by LiveJournal users in languages other than English? What factors favour the

maintenance and spread of these networks? And what are the characteristics of bloggers who occupy bridging positions between languages where they exist?

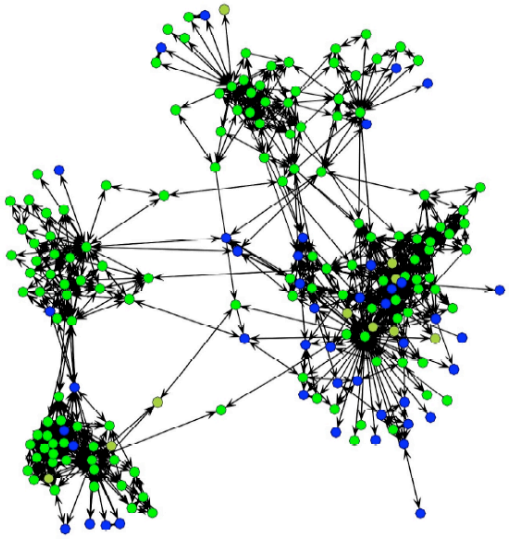


Fig. 2 Example of the Finnish language network on LiveJournal, from Herring *et al*, (2007).

Each of the four language communities was described in terms of size, density and connectedness. Among its many findings, the study concluded that some languages were considerably more persistent than others across the network. This was measured by the continued presence of each language as one spread outwards from the starting sample. Russian was highly persistent in contrast to Japanese, which was only evident in a small fraction of blogs separated by two degrees from the starting sample. This was one of the key factors in the language networks' robustness. English was found to be a much-used *lingua franca*, linking bloggers that otherwise did not share a common language. More specifically, three main patterns of bloggers bridging between languages were identified. The first comprised students of a foreign language who used blogs in different languages to practice their skills and to reach a wider audience through LiveJournal. The second common pattern was that of expatriates bridging between blogs in their native language and the language of their current home. The third pattern included blogs that were comprised mainly of non-verbal content like photographs and audio tracks. It was found that bridging blogs tended to be multilingual and multicultural, often with content of international appeal.

The themes described above in the work of Paolillo and Herring *et al* may be adapted to other research projects. The observations around social position and network robustness are particularly valuable in forming research questions in the context of CMC and language use. The social position of a node may be determined in a number of ways. Firstly, it is measured in terms of its centrality or peripherality in the social network. The most central nodes are those connected to others in the network through the shortest paths. Central nodes are likely to share strong multiplex ties with a relatively high number of other central nodes. As a result they have the quickest and easiest access to information and resources within the network. Peripheral nodes, on the other hand, are connected via longer paths, or may lack a direct path to core nodes altogether. They share relatively fewer and weaker ties to others in the network. Information and resources take longer to reach them, or may not reach them at all. Nodes may exhibit local or global centrality, depending on whether their connectivity is measured in terms of their adjacent group or across the entire network (Scott, 2000). It is a relative measure, and varies according to many factors, including whether or not the direction of links is taken into account.

The reciprocity of links between nodes is another determinant of social position. Reciprocity reflects the hierarchy and power dynamics within a social network. A high indegree signifies popularity and activity within a network, whereas a high out-degree, with relatively few reciprocal links, may signify passivity. A tightly connected group of highly reciprocal links may be classed as a clique. Cliques may exhibit their own norms, values and sub-cultures, and play important roles within social networks, contributing to members' sense of identity and belonging (Scott, 2000). The rate of inward and outward links also points to a node's role within the network (e.g. the 'authorities' and 'hubs' in Herring *et al*, 2005). According to Gibson *et al* (1998: 1), "communities can be viewed as containing a core of central, 'authoritative' pages linked together by 'hub pages'". The relationship between authorities and hubs is a mutually reinforcing one: good hubs point to many good authorities, and good authorities are pointed to by many good hubs.

The role of bridging nodes is an important one. These nodes blur the edges between different groups in the social network. They may facilitate the quicker spread of information and innovation between disparate groups by creating, or shortening, paths between nodes that would otherwise remain distant. Bridging nodes may not

hold any central or authoritative position in the social network, and yet they may play a vital role in the robustness of the network as a whole.

The robustness of a network is a measure of how likely the network is to endure. It is of particular interest in the study of CMC – where communities are often subject to changes in trends and technology; and of sociolinguistics – where a network's language use may be affected over time by changes in its membership. A social network is likely to be robust if it is large and dense with a well-defined core. These are relative rather than absolute measures. For a linguistic network to remain robust, nodes must link to other nodes in the same language. Networks are more likely to be robust if they interact through more than one medium of communication (Haythornthwaite, 2002). If members share a repertoire of ties across multiple media the network is unlikely to disintegrate if one medium breaks down. One key factor contributing to a network's robustness is its persistence. Persistence measures the rate of continued presence of a phenomenon as you spread outwards into the network. The persistence of Russian was reflected positively in the relative size and density of the language's network in Herring *et al* (2005).

D. COMPUTER-MEDIATED DISCOURSE ANALYSIS

1. Computer-mediated Communication

Put simply, CMC is communication that takes place between human beings "via the instrumentality of computers" (Herring, 1996: 1). It is a broad church, encompassing communication through a wide range of electronic media and discourse, most notably across the internet and using the world wide web. Although the term CMC is used by many scholars, it is not an uncontested one with, for example, David Crystal preferring the term 'electronically-mediated communication' to include the increasing use of electronic devices other than computers to communicate. Mobile phones, for example, have made the use of online communication, particularly through social media, possible and popular while on the go.

It is argued in general that the way text is used in discourse online, often in a spontaneous and casual manner, represents a new type of "orality" expressed on screen (December, 1993: online). Crystal uses the term 'Netspeak' to describe in general the characteristics of language use online, claiming it forms a new medium in

its own right, as distinct from written, spoken and signed language as they are from each other. Crystal positions Netspeak on a spectrum between written and spoken language, sharing characteristics of both (Crystal, 2006). We can see this phenomenon online in the many creative ways that users adapt text-based media to carry nuance, humour and emotion by developing unique styles of writing characterised by abbreviations, emoticons and non-standard spellings, for example (Crystal, 2006). This has acted to shrink and blur the differences between live and mediated encounters.

This is a fair description of CMC in general. However, as Crystal shows in his description of language use in blogs, chatrooms, social networking sites (Crystal, 2006) and Twitter (Crystal, 2011), the web is better understood as a number of distinct forms of discourse. Crystal describes them as distinct 'outputs', arguing that the more widely used term 'genre' suggests a "homogeneity which has not yet been established" (Crystal, 2011: 9). This researcher accepts that each web genre described in this study could be further divided into subgenres, exhibiting different language varieties and styles of discourse. For example, the genre of blogging might be subdivided into web journals, filter blogs and knowledge blogs, depending on the type of content shared by the blog authors (see Blood, 2002 for a more detailed description of these blog types). Other genres, like emails, might be described separately depending on the audience and the communication context, with clear distinctions between formal work-related emails and casual emails to friends, for example. However, the term 'genre' is still used in this study to group similar styles of discourse together at a macro-level, according to their loosely shared characteristics. It is inspired by Bauman's (2001: 58) definition of genre as "a constellation of systemically related, co-current formal features and structures" that act as a framework for the production and reception of discourse. Just as we would expect different styles and structures of language use in traditional written genres like letter-writing, diary entries, or telegrams, we might also distinguish between the language style and structure expected in emails, blog entries and Twitter messages.

This study now turns to the work of Susan Herring and Lynn Cherny, and their approaches to defining and describing different web genres.

2. Describing Web Genres

Some core characteristic features and structures should be shared across web genres, even if the styles of discourse that appear there vary between individual sites and users. Distinguishing features include different levels of anonymity and synchronicity; word limits, e.g. maximum 140 characters in a Twitter message; and language styles, e.g. the relative formality of email compared to Facebook messages. To help distinguish between forms of online discourse, Herring (2007) proposes a faceted classification scheme for computer-mediated discourse. These factors inform the users' rationale for using different types of discourse for different occasions, as well as influencing the registers, routines and norms of behaviour that gradually emerge between participants over time on individual websites.

Medium factors	**Situation factors**
Synchronicity	Participation structure
Message transmission (1-way vs. 2-way)	Participant characteristics
Persistence of transcript	Purpose
Size of message buffer	Topic or Theme
Channels of communication	Tone
Anonymous messaging	Activity
Private messaging	Norms
Filtering	Code
Quoting	
Message format	

Table 2. Herring's Medium and Situation factors (2007)

Herring created this scheme primarily as a "focused lens through which to view CMD [computer-mediated discourse] data in order to facilitate linguistic analysis" (Herring, 2007: online). It combines criteria from traditional models of discourse classification to create a list of ten *medium factors*, determined by the sites' technological contexts, and eight *situation factors*, determined by the sites' social contexts. The medium factors describe the influence of computer hardware, software and interfaces on the discourse, and include the websites' levels of synchronicity and anonymity, and message formats. The situation factors describe information about the participants and their shared relationships, and include factors such as the tone, topics

and norms of behaviour (see Table 2). Each of these factors has been observed to affect CMD in some cases, but not all factors will be evident in every website. For example, not all websites allow users to send private messages to other users, or to filter out unwanted messages. These features enable the researcher to clearly distinguish between web genres, and to identify differences between individual websites, using directly comparable factors grounded in academic study. Not only does the classification scheme aid empirical analysis by distinguishing sites of discourse from each other, but it also enables the researcher to group websites by their shared discourse features.

3. Herring's Approach to CMDA

Computer-mediated discourse analysis (CMDA) within networked environments can help describe the dynamics of online communities by observing and coding the types of exchanges taking place between participants. It is a methodology for understanding online behaviour through language-focused content analysis. Herring (2004) summarises the theoretical assumptions that underlie CMDA as follows. Firstly, it can be assumed from established discourse analysis theory that discourse exhibits recurrent patterns. These recurrent patterns may be made consciously or unconsciously by the participants. A basic goal of discourse analysis is to identify and analyse these patterns in an empirical way. Secondly, it can be assumed that discourse involves speaker choices. Thus, as well as analysing linguistic phenomena, discourse analysis must consider the wider social factors affecting speaker choices. A third assumption made specifically in the context of CMC is that, computer-mediated discourse may be shaped by the technological features of CMC genres. In summary then, the basic objective of CMDA is to establish recurrent patterns of discourse among participants, and to analyse them in terms of the linguistic, social and technological contexts within which they are set. In doing this, some basic requirements must be met by the researcher in conducting CMDA. This process is described by Herring (2004) as follows.

- The researcher must pose research questions that are, in principle, answerable. They should ask about empirically-observable phenomena, which are demonstrably present in the data. They should be non-trivial, and should also be motivated by a hypothesis, even if it is just an informal hunch. Herring claims that such questions prove more interesting and more interpretable to the

reader. Finally, they should be sufficiently open-ended that they might lead to unexpected findings.

- The researcher must select methods that address the research questions. Herring stresses that the CMDA approach should not be considered a single method, but rather a set of methods, from which the researcher selects the most appropriate to them, depending on their data and research questions. These methods may include purely quantitative methods – coding and counting chosen criteria; purely qualitative methods – observations of discourse phenomena made, illustrated, and analysed by the researcher in various ways; or a mixture of methods.
- Next, the researcher must apply these methods to a corpus of data sufficient to the scale of the study, and appropriate to the type of phenomena being analysed. CMC can provide a rich source of data for discourse analysis, as many websites produce publicly accessible material as participants comment and respond to each other. Data from more private CMC genres, such as email, instant messaging, and social networking sites, may be much more difficult to source. It is also worth noting that, depending on the levels of anonymity of the site, user characteristics (e.g. gender, age, location) may be impossible to identify reliably, and may hinder the researcher from placing their CMDA within a wider social context.
- Herring lists six different data sampling techniques, and compares their advantages and disadvantages. She concludes that data samples distinguished by *time*, that is, all messages made within a specified time frame, preserve the richest context in which to analyse discourse. If the data is captured from a sufficiently long time period, it is most likely to include coherent threads of discourse as they occur naturally. Striking a balance in setting an appropriate time period for analysis is crucial: too short, and interactions may be truncated; too long, and the data sample may become too large to analyse. Moreover, the researcher must choose a time period that will produce a data sample representative of the typical interactions taking place on the site. To this end, Herring advises supplementing discourse analysis with participant observation on behalf of the researcher. In this way, the researcher becomes more familiar with what constitutes typical day-to-day interaction within the

group under study. They may also witness some interesting interactions taking place outside the data sampling period, which may help inform later analysis.

- Depending on the chosen methods of analysis, the researcher must determine appropriate coding categories to the phenomena under study, and establish their reliability. The criteria for coding categories must be described clearly, so that other researchers returning to the data could in principle identify the codes independently, and test their validity. We have already seen Herring's six criteria for identifying online communities, for example. At this stage the coding can be applied, either manually or with the aid of various computer-aided discourse analysis software. Alternatively, if statistical methods are to be used, the researcher must identify and apply appropriate statistical tests.
- Finally, the findings must be interpreted and published responsibly. These findings should refer back to the original research questions from which the methods of analysis were established.

4. Linguistic Analysis in the Work of Lynn Cherny

In *Conversation and Community* (1999), Lynn Cherny conducted an ethnographic study of a community of MUD users. In MUDs, users express their thoughts, utterances, emotions, actions and environments through text on screen. Taking on the role of a participant-observer within the community, she analysed the relationships formed between users through their on-screen discourse. Cherny firstly conducted a linguistic analysis of interactions typical of members' conversations and described various features of the MUD register including: syntactic and morphological variation; abbreviations and shortenings; routines; and play with modality.

- Syntactic and morphological variation: Studies into various web genres such as MUDs (Cherny, 1999), chatrooms (Paolillo, 2001), and micro-blogging sites (Crystal, 2011), commonly find syntactic and morphological variations, primarily made to economise on typing effort, mimic spoken language features or for creative expression (Herring, 2001). In Cherny's study these included word contractions, non-standard verb formations, and word deletions. In an on-screen, text-based environment some non-standard features give clues to how participants use the language in face-to-face verbal discourse.
- Abbreviations and shortenings: These are short abbreviations used to express

common phrases or emotions. Abbreviations are commonly made in CMC to speed up communication or to conform to message buffers. Examples of abbreviations identified in Cherny's MUD and common to CMC in general include: LOL (Laugh Out Loud), btw (by the way) and WTF (What The Fuck?).

- Routines: Routines are standardised patterns of speech that are frequently uttered in a speech community. They act as group rituals that give a feeling of membership to a community and a "sense of confidence and behavioural certainty" to their members (Cherny, 1999: 95). In this way they are important in fostering a sense of belonging among group members. Cherny identified language routines in the use of ritualised responses to specific types of questions, statements and roll calls.
- Play with modality: This describes the written features that denote the mode, mood, or manner in which posts and comments are made. In the absence of audiovisual cues such as posture, eye contact and volume change, web users use innovative graphical variations to reflect the emotion, intonation, stress, speed, rhythm, pause, and tone of voice that occur in face-to-face communication (Crystal, 2006). These features go some way in preventing readers from misinterpreting the tone of messages from how the writer intended. Perhaps the most striking of these features is the emoticon, or smiley face, which is prevalent across many web genres.

This study will use Cherny's categories to describe the discourse shared between core members of the networks under study. This is done despite the fact that Cherny's work was produced over ten years before this study (a long time in internet terms) on a web genre that appears out-dated today. Each researcher must choose their own criteria and codes for discourse analysis, based on the research questions they seek to answer. To illustrate the wide range of issues that might be explored through computer-mediated discourse analysis, and the phenomena that might be analysed and coded in the process, Herring created the matrix of approaches to CMDA in Table 3. It shows the spectrum of phenomena that might be analysed and the large number of analytical methods that might be employed. This study seeks to analyse some of the structural phenomena of Irish language discourse online, to show how users express

themselves online in sometimes new and innovative ways. In the process, it seeks to analyse some of the social dynamics of users, to gain a better understanding through analysing their discourse of how the groups behave as online communities. Despite Cherny's work focusing specifically on one MUD community, she sought to explore the same issues of discourse structure and social dynamics.

Levels	Issues	Phenomena	Methods
Structure	Orality; formality; efficiency; expressivity; complexity; genre characteristics, etc.	Typography, orthography, morphology, syntax, discourse schemata, formatting conventions, etc.	Structural /Descriptive Linguistics, Text Analysis, Corpus Lingustics, Stylistics
Meaning	What is intended What is communicated What is accomplished	Meaning of words, utterances (speech acts), exchanges, etc.	Semantics, pragmatics
Interaction management	Interactivity; timing; coherence; repair; interaction as co-constructed, etc.	Turns, sequences, exchanges, threads, etc.	Conversation analysis, ethnomethodology
Social phenomena	Social dynamics; power; influence; identity; community; cultural differences, etc.	Linguistic expressions of status, conflict, negotiation, face-management, play; discourse styles/lects, etc.	Interactional Sociolinguistics, Critical Discourse Analysis, Ethnography of Communication
Multimodal communication	Mode effects; cross-mode coherence; reference and address management; generation and spread of graphical meaning units; media co-activity, etc.	Mode choice; text-in-image; image quotes; spatial and temporal positionality and deixis; animation, etc.	Social semiotics; visual content analysis; film studies (?)

Table 3. Five levels of CMDA (adapted from Herring, 2013)

The researcher might have chosen to base its discourse analysis on the more recent approach used by David Crystal in his 2011 guide to internet linguistics. In this book, Crystal uses Twitter as a 'microexample' of online discourse. Extracting 146 tweets on the theme of 'language', he meticulously analyses their content, grammar and pragmatics. In doing so, he describes many features of Twitter including shortening techniques, contractions, logograms, abbreviations, spacing, ellipsis and punctuation. The researcher found many of these features in Irish language discourse on Twitter and other social media, and Crystal's work is referenced frequently later in this study. However, the objective of this research is not to 'tick off' a list of features of English language CMD that also appear in Irish language online discourse. Rather, it aims to

paint a portrait of how the language is used and, importantly, describe how the groups under study behave like online communities. Cherny's headings create a loose framework through which different features might be grouped, without being overly-prescriptive about what was expected to appear. Cherny's headings also align with the criteria for online communities proposed by Herring and Baym, and used later in this study.

Contemporary studies into CMD tend to focus on one phenomenon in-depth, using samples of discourse specifically chosen from the web because they contain the phenomenon under study. For example, in three volumes of the online journal *language@internet* (2010-2012) the following articles feature focused analysis of individual discourse phenomena in specific web genres, including: the use of metaphor in themed YouTube video threads (Pihlaja, 2011); codeswitching on Usenet and IRC (Paolillo, 2011); cues for participant alignment in email discourse (Georgakopoulou, 2011); rapid language change in email discourse (Rowe, 2011); representations of consonants in Jordanian online chat (Jarbou & al-Share, 2012); discourse structures in instant messaging (Baron, 2010); quoting in email and newsgroups (Severinson Eklundh, 2010); and turn-taking in a synchronous CMC system (Anderson, Beard & Walther, 2010). Of course, there is only space in journal articles to focus on one or a few specific features, and many of the articles listed above build on an established corpus of research already published on CMDA in their languages. This study, on the other hand, seeks to paint a general portrait of Irish language discourse in the blogosphere, on Twitter and on Facebook, on which there is very little written, and to this end it needed to analyse a broad range of criteria.

As a first step in describing the discourse in the groups in this study, Cherny's criteria cover a wide spectrum of discourse features that are of interest to the researcher, in not only describing some of the innovative linguistic features but also the ways in which language style and routines act to bond members of networks. Her work took place in a monolingual English language MUD community, however, and in the context of Irish language CMC there were two further linguistic features that, it was felt, merited analysis: code switching and dialectical variation. These themes will be introduced later in the study. It is hoped that combining these two themes with Cherny's criteria will result in a reasonably in-depth introduction to how the Irish language is used interactively online today.

After considering the above literature on computer-mediated communication and the various approaches researchers have taken to analysing online discourse, the researcher now turns to the specific aims of this research project in the context of the Irish language online. These are outlined below, and clearly defined in a list of research questions.

III RESEARCH QUESTIONS

This study began with the hypothesis that the Irish language is being used in small networks of users online, and that some of these networks comprise online communities of regular users interacting in socially meaningful ways. In determining where and how the Irish language is used online, and whether online communities of Irish language users do indeed exist, this study posed a number of research questions. These questions began broadly with an overview of the vitality of Irish language use online, gradually funneling down to the language's use in specific online groups. Firstly, it established the level of online activity in Irish in general, with reference to other minority languages, in response to the question:

RQ1.1: How does Irish language activity online compare with that of other European minority languages?

This involved conducting an audit of Irish language use online through statistics and user data that could be directly compared to other languages. This offered a relative measure of online ethnolinguistic vitality for the Irish language, and shed light on some of the contexts within which the language is used.

The focus then turned to identifying those web genres in which the Irish language is currently being used interactively online. The objective was to introduce the reader to the landscape of websites that are participated in publicly by significant numbers of Irish language users, and to identify research sites for later discourse analysis:

RQ1.2: Where do significant numbers of people interact publicly in the Irish language online?

This question was answered by analysing popular social networking sites, discussion forums, chat rooms, blogs and websites for evidence of Irish language use. It was important that these sites were publicly accessible, so that the types of interaction taking place could be described. Therefore, private networks of Irish users facilitated through social networking sites, e-mail, multiplayer online gaming, instant messaging, etc., where users must acquire permission to observe what's going on, remained outside the scope of this study. The sites with the most significant examples of Irish language interaction were compared with reference to Herring's faceted classification scheme for online discourse in terms of: their design; levels of anonymity, synchronicity, accessibility, and exclusivity; and the types of users they attracted, languages they used and messages they shared. In this way, the types of computer-mediated discourse currently most attractive to Irish language users were identified.

The study then selected three of the most significant groups of Irish language users online, from the web genres identified above, as individual case studies. The activities of each group were observed over a period of time and interactions between members were recorded. Social network analysis techniques were used to answer the following question:

RQ2.1 How large is each network, and within each network how interconnected are their members?

In answering this question, the study described the connections between members in each group in graph terms and created computer-generated visualisations of each network using observable relational data from a specific period of study. The visualisations helped establish the social structure of each network and each member's position within it. In particular they helped identify core and peripheral members within each network. Once this was established, the focus turned to describing the role of language, geography and other distinguishing features in the relative position and participation of group members. The study sought to answer:

RQ2.2: Can social position in each network be explained by geographic location, language use, subject matter, or length of activity?

In answering this question, this study determined where Gaeltacht dwellers/non-Gaeltacht dwellers, monolingual Irish users/infrequent Irish users, newcomers/established users, and individuals interested in various topics were positioned in the social network in relation to each other, and how well they were assimilated into the core. Further social network analysis sought to answer the question:

RQ2.3: How robust is each network?

This question stemmed from an interest in how well networks of Irish users are maintained within a greater multilingual worldwide web. Network robustness was described in terms of tie density, multiplexity and strength, and with reference to the 2007 study of blogs by Herring *et al* (2007), in which robustness is defined in both structural and linguistic terms.

Linguistic analysis was conducted on a corpus of content from core members of each network, as identified through social network analysis. This sought to answer the question:

RQ3: How has the Irish language been adapted by core members of each network in a text-based computer-mediated environment?

Specifically, the study sought to describe the register of language use in each group according to the categories outlined in Cherny's (1999) study of MUDs. Moreover, the researcher established what role, if any, dialect plays in promoting or obstructing group members from forming connections.

Finally, computer-mediated discourse analysis then sought to answer the question:

RQ4: How is each network characterised as an online community?

This was answered with reference to Herring's (2004) criteria for online communities. SNA revealed the groups' participation structures and some aspects of group roles

and hierarchies. Content from core members of each network was analysed and coded according to how members expressed shared culture and norms; how they shared solidarity and support; how they resolved conflict between group members; how they expressed an awareness of their group being different from others; and if and how roles and hierarchy were manifested within the groups.

IV METHODOLOGY

Each stage of research required a separate methodological approach tailored to the research question(s) being answered. These are outlined below.

1.1 Comparing Irish language activity online with that of other European minority languages

No statistics exist for the number or proportion of Irish language web pages online. The Irish language, like most other minority languages, is not included in any of the reports or studies cited by Gerrand (2007) or Cunliffe (2007) (though one is referenced in Deere (2011)). Rather than producing an estimate of Irish language activity in isolation, it was decided to compare the vitality of Irish on the web with those of other minority languages. The Irish language will always comprise just a fraction of overall online activity, paling into insignificance compared to global languages like English, Chinese or Spanish. However, a like-for-like comparison with other minority languages, in varying degrees of endangerment, would give a more valuable insight into the relative vitality of the language online. Moreover, rather than gauge the strength of the language according to just one of the approaches described by Gerrand (2007), each with its own pros and cons; the researcher combined aspects of user profile, web presence *and* user activity to give a relative measure of the languages' vitality online. The results were combined to rate each language according to UNESCO's grades of response to new domains and media, according to a spectrum of activity from dynamic to inactive. This gave a wider context to the rest of the study as it proceeded to fix its focus solely on the Irish language.

Five languages were chosen for comparison: Irish, Scots Gaelic, Welsh, Basque and West Frisian. All five languages represent minority languages[22] indigenous to western European countries – three of which come from the Celtic family of languages. The countries with the largest speaker populations – Republic of Ireland, United Kingdom, Spain, and the Netherlands – are all developed countries in the EU in which access to computers and the internet is widely available and social media use on various websites is popular. In fact, they were the five most commonly used European minority languages on Twitter according to data gathered by the website

[22] As mentioned previously, 'minority language' is a contested term. Although there is a big difference in the numbers of speakers of each of the languages listed here, and their regional vitality, each language *is* a minority language within its own country and they all appear on UNESCO's list of endangered languages.

Indigenous Tweets.[23] The researcher had at least some working knowledge of Irish, Scots Gaelic and Welsh.

Figures from reports and websites where the five languages were directly compared like-for-like were gathered for analysis. These were grouped as follows:

- User profile: The potential user profile for each language was deduced by combining speaker population data with statistics on internet penetration in the countries with the largest speaker populations. Population data was compared from three sources: Ethnologue, UNESCO and studies cited by Wikipedia. Internet penetration statistics were sourced from figures compiled by Eurostat.
- Web presence: A relative measure of web presence in the five languages was made through searches on three of the most popular web search engines internationally: Google, Yahoo! and Bing. Admittedly, the chosen search terms and the ways in which the searches were conducted (e.g. with/without inverted commas, Boolean commands, alternative spellings, etc.) would greatly influence the results. In particular, it was expected that the highly inflected languages – Irish, Scots Gaelic, Welsh and Basque – would be under-represented in their search results. A further comparison of content was made between the number of Wikipedia articles and Wiktionary entries in each language. Furthermore, Deere's (2011) study into minority language presence on Wikipedia was referenced. It was hoped that because of the wide international appeal of Wikipedia in particular, and the fact that editions of the website in the five chosen languages are well established, that the figures could be used as a guide to relative web presence in general.
- User activity: Where possible, the numbers of individuals using some of the most popular social media sites internationally were directly compared. These included registered and active users on Wikipedia, and the number of Twitter accounts and blogs using the languages as identified by Indigenous Tweets and Indigenous Blogs.[24] Further analysis of the linguistic activity of the most prolific Twitter users in the chosen languages was carried out to compare how frequently users tweeted in each language.

[23] Data sourced from Indigenous Tweets (http://www.indigenoustweets.com) in July 2012.
[24] Indigenous Tweets (http://www.indigenoustweets.com), Indigenous Blogs (http://www.indigenoustweets.com/blogs).

Combining these samples of user profile, web presence and user activity, the researcher graded the languages according to UNESCO's methodology for assessing languages' response to new domains and media. This would describe the general vitality of the Irish language in an online context, in direct comparison to the other chosen languages, on a spectrum from 'dynamic' to 'inactive'.

1.2 Identifying the most popular sites of regular Irish language activity online

The rest of this study is concerned with the Irish language only and those websites where it is used interactively. The criteria by which these websites were searched for, identified and analysed needed to be carefully considered. Most CMC researchers begin their studies with a narrowly defined target of their chosen websites or web pages of interest. Examples include Trammel *et al*'s (2006) random sample of 500 blogs from Polish blog hosting service blog.pl; Cherny's (1999) participant-observer study of one MUD community called ElseMOO; Lysloff's (2003) online ethnography of the "mod scene", communicating and sharing music via "chat" systems and email; or McKenna & Bargh's (1998) content analysis of several newsgroups associated with marginal identities. Each of these studies gives a picture of the wider context and history within which these groups are set. However, their starting points for analysis are very narrowly defined, and conclusions are made within the small worlds that exist within their groups' boundaries. In contrast, the current study began with a much broader analysis, or inventory, of those sites where the Irish language is currently used interactively, only later refining the research to more narrowly defined groups. The inventory needed to cover a wide sample of Irish language websites, using search criteria uniquely tailored to where the language was most likely to be found.

Conducting an inventory of Irish language websites

To set the research within the wider landscape of Irish language use on the web, this study created an inventory of Irish language content websites drawn from a number of separate online searches. By approaching the search from a number of angles adapted from the methodologies of Kelly-Holmes (2006) and Wright (2006), it was hoped that a reasonably representative sample of popular Irish language websites would be found. The online search was conducted as follows.

Firstly, the search explored the websites listed in two Irish language *portal* sites. Portals are "thematic indexes to the Internet" (Cunliffe, 2007: 134): individual

websites that provide links to a set of websites that form a related community or theme. Kelly-Holmes (2006) identified two important portals dedicated to promoting the Irish language: Gaelport,[25] and the Irish language section of the *Sabhal Mòr Ostaig* (SMO) website.[26] Gaelport is a site produced by the Irish language organisation *Comhdháil Náisiúnta na Gaeilge*. Each of the organisations listed in its 'Member Organisations' were added to the inventory. The SMO website is a particularly prolific portal site, produced by a Scots Gaelic university on the Isle of Skye, Scotland, with "over 700 classified links to pages in and about Irish Gaelic" according to the site. The links in these portal sites were supplemented by two other sources: the URLs listed in Ní Chartúir's book *The Irish Language: An Overview and Guide* (2002), and those listed in the Irish language section of the Irish Government's Department of Community, Equality and Gaeltacht Affairs website.[27] These were chosen as authoritative sources of information about the Irish language.

The second approach involved Google searches using simple Irish language terms that *might* point to clusters of Irish language users. Google is by far the most popular search engine in Ireland[28] and is considered by Kelly-Holmes to be particularly suitable in searching for websites due to its popularity and its use of "ranking, relevance, and links into and out of the site to list results" (Kelly-Holmes, 2006: 221). Specifically, this study used the Irish language version of the site available at: http://www.google.ie. The objective of Kelly-Holmes' study was to see the range of domains in which Irish is represented online, and her search terms considered both what she called 'typical' and 'non-typical' domains for the Irish language (ibid.: 218). This current study was specifically interested in exploring sites where Irish was used interactively and the search terms reflected this particular focus. The top 50 results for the six chosen search terms were analysed. Firstly, in line with Wright's search for minority language websites, this study searched for the Irish terms for the language itself, "Gaeilge" (Irish language); as well as the adjective related to the language, "Gaelach" (Irish). These were supplemented with four other common terms related to interactivity: "teagmháil" (contact), "pobal" (community), "cairde" (friends), and "plé" (discuss) as chosen by the researcher. These terms were chosen for their likelihood to appear in interactive Irish language content. Irish is a highly-inflected

[25] http://www.gaelport.com
[26] http://www.smo.uhi.ac.uk/ gaeilge/gaeilge.html
[27] http://www.pobail.ie/ie/AnGhaeilge/Naisc
[28] According to Alexa, Top Sites in Ireland: http://www.alexa.com/topsites/countries/IE.

language and, under certain grammatical conditions, nouns and adjectives in the Irish language are frequently subjected to morphological change associated with lenition, initial consonant mutation, and declension. The morphological variants have the same meaning as the standard dictionary definitions, but reflect the different spellings of the words as they may appear in the context of other words around them. The morphological variants of the six search terms were sourced from the online Irish language dictionary *An Foclóir Beag,*[29] deemed the best of the online Irish dictionaries for this specific task, and were included in the search. Using simple Boolean commands, the searches, therefore, consisted of the following search terms: "Gaeilge OR Ghaeilge OR nGaeilge", "Gaelach OR Ghaelach OR nGaelach OR Gaelaigh OR Ghaelaigh OR Gaelacha OR Ghaelacha OR Gaelaí OR Ghaelaí", "teagmháil OR theagmháil OR dteagmháil OR teagmhála OR theagmhála", "pobal OR phobal OR bpobal OR pobail OR phobail", "cairde OR chairde OR gcairde", and "plé OR phlé OR bplé". It is important that any online searches are sensitive to the unique traits of the language under study. A similar approach to morphological variants was necessary in Honeycutt and Cunliffe's (2010) study of Welsh language groups on Facebook, where they carried out searches using the terms "cymraeg" (Welsh) and "cymru" (Wales) as well as their morphological variants "gymraeg" and "nghymru".

In addition to these generic search terms, the researcher carried out a Google search using the standard Irish language dictionary definitions of seven popular interactive web genres. Seven of the terms (there are two different terms for 'bulletin board') were sourced from Focal.ie,[30] deemed the best online dictionary for this specific task. Again, the top 50 results were explored for examples of websites with Irish language content. The search terms and their translations were as follows: "fóram plé" (discussion forum), "seomra comhrá" (chat room), "blag" (blog), "clár fógraí" (bulletin board), "clár feasachán" (bulletin board), "láithreán líonraithe shóisialta" (social networking site), "grúpa nuachta" (newsgroup), all sourced from Focal.ie, and "clár plé" (discussion board), which was observed in use on other Irish language websites. It was not known how widely used these phrases are, although it was conjectured that the more difficult translations, in particular "clár feasachán" and "láithreán líonraithe shóisialta", were unlikely to be used popularly. The

[29] http://www.csis.ul.ie/focloir
[30] http://www.focal.ie

morphological variants were not included this time. Whereas the previous Google search was looking for generic terms found in Irish language content, it was hoped that the second set of terms would appear in tabs and sitemaps. These are most likely to use standard dictionary definitions of the terms alone and without any morphological variation.

The fourth approach to discovering Irish language websites explored some of the most popular interactive websites nationally and globally – those which were likely to attract Irish participants. The top 150 most visited websites in Ireland as listed by Alexa[31] were explored for Irish language content. In this way, popular websites such as Facebook, Twitter, MySpace, Wikipedia, YouTube, Blogger.com and Flickr were analysed for evidence of Irish language use among their members. The process of individually interrogating each of these websites for Irish language content varied according to the websites' design. If the website had a sitemap it was explored for references to Irish language content. As the most popular websites in Ireland, it was expected that they would be well designed and easy to navigate, which would help greatly in the search.

It was assumed that by combining these four search strategies, the inventory would represent a wide range of the most popular websites with Irish language content.

Weeding the inventory

As the researcher visited the sites in the inventory, they were interrogated and filtered in the following ways. Firstly, any websites which had been deleted or no longer pointed to Irish language content were removed from the sample.

Secondly, sites using the Irish language in only tokenistic ways were removed. In her study into Irish language websites, Kelly-Holmes coded the websites she visited according to the amount of Irish language used. One of her categories distinguished those websites that she described as being linguistically "English with only symbolic use of a minimal number of Irish words" (Kelly-Holmes, 2006: 223). She accepted that this distinction was open to criticism, but argued that it was necessary to establish whether the use of Irish went beyond symbolic or tokenistic switching or, in her words, "linguistic fetish" (ibid.: 223). Examples that appeared in her searches included a students' union shop in University College Cork called *An Siopa* (the shop)

[31] http://www.alexa.com/topsites/countries/IE, accessed 17/11/2011

on their website but with the rest of the information in English; and the website of traditional Irish musicians and dancers called *Ceol Chiarraí* (Kerry music) which, apart from the group's name, was composed entirely in English. Any websites observed using Irish in this way were removed from the inventory.

Finally, the sites were categorised according to a scheme adapted from the nine categories of domain used by Kelly-Holmes. It was considered necessary to adapt the categories, rather than apply them directly, for three reasons. Firstly, it was difficult to identify the sources of many websites. Secondly, there was an issue of ambiguity around those websites that did not fit neatly into Kelly-Holmes' categories. She herself felt it necessary to categorise some of her results in more than one domain. Thirdly, it was felt that Kelly-Holmes' categorisation lacked an appropriate category for the many community-based groups whose main objective was to promote the use of Irish in the community. These groups most often fit somewhere between the level of official and individual. The four primary categories used by this study described the websites according to whether they were produced by:

- *OF*, *Official organisations*: The websites of Government – national and local, as well as state-funded organisations, including educational institutions and media organisations, like RTÉ and BBC, which are subject to statutory rules.
- *CO, Private Commercial organisations*: The websites of profit-orientated organisations. This category included the language schools, online language courses, and translation services that Kelly-Holmes describes separately as 'language industry'.
- *TS, Third Sector organisations*: The websites of community-based, voluntary, or charitable groups, to include many of the sporting, arts and entertainment organisations categorised separately by Kelly-Holmes. These are removed from government although they may receive some state funding through various grants.
- *PI, Personal Interest*: The personal homepages of individuals, as well as websites about shared interests or hobbies contributed to by more than one person.

Each of the above categories described the web offerings of organisations and individuals with different objectives in producing online content in Irish. These

objectives may variously involve statutory obligation, commercial advantage, language cause or individual interest. This system described the sources of the websites in a clear and unambiguous way that, it was hoped, would be easily understood by the reader.

This categorisation is not an exact science but, rather, was left to the researcher's own judgment. However, the categories were useful in gauging the breadth of the sample and identifying any bias within the searches. It should be noted that Wright's (2006) study into six regional and minority languages in France, Italy and the Netherlands described her search results in similar terms to Kelly-Holmes, using the categories: *official, associations of civil society, education, business* and *individuals.*

Testing the inventory

It was important to test the base sample in some way to see if it represented a significantly wide range of Irish language online content. To this end, two further steps were built into the methodology to examine the sample's validity.

The number of sites in each category was quantified and compared to test the breadth of the sample. If one or more categories were deemed to be poorly represented in comparison to others, it would require consideration of whether the searches were biased towards particular types of websites and whether further searches needed to be carried out. Here, two predictions were made according to Kelly-Holmes' (2006) and Wright's (2006) studies. Firstly, Kelly-Holmes (2006: 217) found that "the 'official' sector and other sectors closely related to language policy and planning" were the main providers of monolingual Irish language material online. We might expect, therefore, the *OF, Official* category to be overly represented. Secondly, both Kelly-Holmes and Wright note in their research that the commercial sectors are underrepresented online in their chosen minority languages. It was likely, therefore, that the *CO, Private Commercial* category would be poorly represented across the sample.

The researcher had built up a short list of Irish language websites from six months of observing Irish language use online. The URLs of those websites had been collected by the researcher as he casually browsed Irish language websites, language resources, blogs, Twitter accounts, and Facebook groups. The base sample was tested to see if it included three websites selected from the researcher's list. These were: the

Daltaí na Gaeilge discussion board,[32] the *People's Republic of Cork* Irish language forum,[33] and the BBC's Irish language web page.[34] These three websites were chosen because they represented regularly updated websites, with interactive elements, at a national and regional level. The reader is reminded that not all websites are found through active searches. The web encourages users to 'surf' from site to site in an unstructured way, through its network of hypertext links, stumbling upon websites as they go. It was hoped that by testing the sample to see if it included these three websites stumbled upon by the researcher, and representative of the types of interactive websites of interest to this study, the success of the searches described above could be validated.

Identifying the most significant sites of interaction

It was assumed that the inventory included the most significant Irish language content websites. However, in order to merit analysis as sites of interaction they had to satisfy three basic criteria, defined as follows.

1. They must be *publicly accessible*. The interactions taking place must be freely observable, without overly restrictive access barriers or charges, or the need to have prior knowledge of existing members. Where simple log-ins were required the researcher created an anonymous account. It is important to reiterate that this study included web-based interaction only. Groups who maintained contact via other internet channels, e.g. email, instant messaging, gaming, etc., remained outside of the scope of this study.
2. Many websites remain online long after they've ceased to be used. In web terms these are called 'dead' websites. This study was only concerned with sources of *regularly updated Irish language content*. It would be too prescriptive to set targets for the frequency of interaction, as websites attract different rates of communication depending on a number of medium and situation factors. However, the patterns of use had to point to a history of activity up to the present day. Multilingual websites were considered, but where the Irish language was used it must have made up at least half of all interactions.

[32] http://www.daltai.com
[33] http://www.peoplesrepublicofcork.com/~peoplesr/forums
[34] http://www.bbc.co.uk/irish

3. They must involve *significant numbers of repeat users*. It must be observable from the usernames and/or content that the site attracted repeat visitors who had experience interacting with each other. Again, this was a relative rather than an absolute measure. It was a useful gauge in identifying those sites where relationships were likely to form between users over time through their interactions.

Describing the most popular web genres

Once the three search criteria were applied to the inventory, the refined sample consisted of a much shorter list of websites. These were grouped according to their web genre, and described according to Herring's (2007) faceted classification scheme. This was followed by a description of the key medium and situation factors that defined the most popular forms of discourse for Irish web users. From the refined list of popular websites, the three websites/genres with the most active groups of users were chosen for further study using social network analysis and computer-mediated discourse analysis techniques. These were: the Irish language blogosphere, Twitter, and the Facebook group *Gaeilge Amháin* (meaning 'Irish Only').

2.1.1 Constructing social networks for each group

The second phase of this study applied social network analysis to the three individual clusters of Irish bloggers, Irish Twitter users and members of the Facebook group *Gaeilge Amháin (GA)*. The term 'clusters' is used here to describe groups of people who could be observed interacting either within or between websites. These research sites each satisfied the three earlier criteria for analysis as follows. Firstly, all of the active blogs, the vast majority of Twitter accounts, and the activity between *GA* members were publicly accessible, without log-ins or subscriptions necessary to read the content. Secondly, although activity levels varied widely between the three groups, each represented a source of regularly updated Irish language content. One of the defining features of all three is that content is listed in reverse-chronological order with the most recent content at the top, making it easy to find fresh material. Messages and responses posted in the three research sites were time and date stamped, making it possible to identify content within specific study periods. Moreover, messages and comments posted to blogs and Facebook groups remained online indefinitely, creating a searchable archive of content from each research site

(using computer software approximately 3,000 of the most recent tweets could be downloaded from Twitter accounts). This made it possible to reference content from the past and to build up a more longitudinal picture of user activity. Thirdly, the content in each group could be attributed to specific users through their unique usernames, and pointed to frequent interaction back-and-forth between some participants, forming what Herring *et al* (2005) called 'conversations' (in the context of blogging). All of these factors pointed to regularly updated, interactive clusters of Irish language users.

Each of the research sites was particularly suited to social network analysis because of the different ways in which users could be observed interacting and linking to others. This created a source of relational data through which social networks could be constructed and visualised. Participants could be observed creating both directed and undirected links to others, revealing patterns of reciprocity between nodes in each network. Moreover, each website/genre afforded its participants the chance to construct their online identities in many creative ways. Not only was their personality expressed in the messages and comments they wrote, but also in the ways they presented their blogs, Twitter accounts and individual Facebook pages and the personal information they chose to share. The ways in which the Irish language was creatively used in each research site, across a spectrum of formality, tone, register and subject matter, represented a rich source of content unique to each web genre. For all of these reasons the clusters of active users identified in the three research sites were singled out as vibrant online Irish language groups worthy of further study.

Setting boundaries for each social network

Each study of CMC must set its own threshold for what constitutes active participation. In making this decision the researcher takes medium factors and situational factors into account, as well as the size of the community and the type of members, communications or relationships of particular interest. As Butts (2008: 17) states in his methodological introduction to social network analysis, "a network is bounded by the set of entities on which it is defined". These entities must be defined in a "substantively appropriate manner" so as to avoid basing the inclusion/exclusion of nodes and ties on mere methodological convenience (ibid.: 17). Thus, the boundaries of the networks in this study were set by specific criteria, chosen by the researcher as relevant to the interests of this study, that were meaningful *vis-à-vis* the

specific conventions of blogging, Twitter and Facebook. Each group was imagined via both methodologically defined and relationally defined boundaries (Butts, 2008). The methodologically defined boundaries refer to the research samples chosen, and the various thresholds for language use and activity. The relationally defined boundaries refer to the types of interactions and connections between individuals through which nodes were connected within their networks. The process is outlined below to add empirical clarity to the types of users and relationships included in the analysis, to aid the reader in interpreting any social network data or visualisations shown later in the study, and to set a benchmark by which further research might be constructed.

Sourcing research samples

The Irish language blogosphere (ILB): In the absence of any specific blog aggregator grouping Irish language blogs together in one website[35] – such as the *Blogiadur* that formed the basis of Cunliffe & Honeycutt's 2008 Welsh language blog study – potential members of the ILB had to be found through a combination of methods. The research began with a twelve-month period of participant-observation. Having created a blog (*Faoi Cheilt*[36] meaning 'hidden') the researcher commenced publishing regular posts, approximately ten per month, on various topics and in the Irish language. During this time other Irish language blogs were found through internet searches, searches in popular blog hosting services, and from lists of blogs in traditional media, e.g. the newspaper *Gaelscéal,* and Ní Chartúir's (2002) book about the Irish language. The results of these searches were used to create a blogroll on the researcher's blog. It is deemed unnecessary to describe these initial searches in greater detail as they were carried out in an unordered manner and the more significant task of expanding the list of blogs occurred later in the research. The researcher read and commented on the blogs in his blogroll, on a sporadic and uneven basis. Where possible, in the case of those hosted by the blog hosting service Blogger, the researcher became a 'follower' of other Irish language blogs. This meant that the

[35] An aggregator of Irish language blogs hosted by NetVibes was discovered at the end of the period of participant-observation. However, the number of blogs included was less than half of the blogs that had already been found through other methods. It offered only a very limited picture of the Irish language blogosphere. The Indigenous Blogs website began listing active Irish language blogs in September 2011, six months after the blog data capture period. However, only one additional blog was found to have posted during the data capture period, and this did not attract any comments or interactions from other users. This blog was added to the network, albeit as an isolated node, at a later stage.

[36] http://faoicheilt.blogspot.com

researcher's avatar[37] appeared publicly on other Irish language blogs in lists of 'followers'.[38] Significantly, the blogroll was expanded over time as new blogs were discovered through hypertext links embedded in blogs the researcher visited, or as previously unidentified bloggers commented on the researcher's blog. By the end of the period of participant-observation the researcher had compiled a blogroll of approximately 40 blogs.

The bloggers and commenters that comprise the ILB in this study remained unaware of the researcher's role as participant-observer before and during the data capture period, without exception. However, a short profile displayed prominently on the front page of the researcher's blog read: *Is mac léinn mé, in Ollscoil Chaerdydd. Táim ag staidéar PhD sa teangeolaíocht, go háirithe conas a chruthaítear pobail ar an idirlíon* (meaning "I am a student, in Cardiff University. I'm studying a PhD in linguistics, specifically how communities are created on the internet.")

Irish Language Twittersphere (ILT): This study used as its source the Twitter aggregator Indigenous Tweets created and run by Prof. Kevin Scannell of Saint Louis University, Missouri. His website collates the Twitter activities of over 100 minority languages, creoles and dialects (see Scannell, 2011). Using a web crawler (Scannell, 2007), it identifies new messages posted to Twitter in these languages and then aggregates the accounts that tweeted them in one list for each language. It also provides additional information about the individuals' activity and the number of messages they have posted in the selected language.

By the data capture period, Indigenous Tweets had identified approximately 2,700 individual accounts tweeting in Irish. However, the site only publicly lists the 500 most prolific Twitter accounts in each language – in the case of Irish those accounts with 33 or more Irish language tweets. Members of what would constitute the ILT in this study were sourced from within this sample. This was a convenient method for compiling a research sample that suited the scale of this study, albeit with some disadvantages. Firstly, it relied on external computer technology for identifying and coding Irish language Twitter accounts, which could neither be controlled nor verified by the researcher. That said, Prof. Scannell is an expert in applying computer technology to linguistic analysis and his research in the field has been widely

[37] An avatar is a graphical representation of a computer user's character – in this case a close-up photograph of the researcher.
[38] This process is described further below.

published.[39] Secondly, there was a threshold of 500 Twitter accounts, excluding from analysis the less prolific Irish language Twitter users (i.e. those who had tweeted fewer than 32 times in Irish since Indigenous Tweets began collecting data on Irish Twitter users). However, the disadvantages were outweighed by the benefits of sourcing the entire research sample from one site and the availability of comparable usage data collected over time. Considering the number of Twitter accounts found tweeting in Irish by Indigenous Tweets, identifying and coding Irish Twitter accounts by hand would have been prohibitively time-consuming. Furthermore, future research into Irish language Twitter activity will be able to draw on the same source of data and make direct comparisons back to this study's sample.

Gaeilge Amháin Facebook group (*GA*): *GA* is the most popular Irish language Facebook group. It is an open group with over 1,600 members in August 2012. Created by Irish language enthusiast Ailéin Ó Clúmháin in 2011, the aim of the group is to encourage Facebook users to communicate with each other through Irish only. The majority of members' personal Facebook accounts were private, and the interactions taking place on individuals' homepages, or walls, remained beyond the reach of the researcher. However, all of the interactions that take place between members on the group's wall are accessible to the public, with or without group membership. These interactions, and the individual members participating in them, would become this study's research sample.

In their study of Welsh language use on Facebook, Honeycutt and Cunliffe (2010) developed an alternative methodology that built a social network of Related Groups popular with Welsh language users. Related Groups is a feature on Facebook whereby a list of groups that have the most members in common are displayed to the viewer, revealing links between clusters of users across Facebook. These groups, and not their members, constituted the nodes in Honeycutt and Cunliffe's social network visualisations. Since the focus of the present study is on individual web users and their connections with other individuals, it was decided to construct a new methodology involving the active members of the *GA* group alone.

[39] For a list of publications and projects, see Prof Scannell's webpage: http://borel.slu.edu/nlp.html

Setting data capture period

ILB: A three-month period of data capture was chosen from January 1st to March 31st 2011. All posts and comments made during this time were analysed. It was hoped that a three-month period would be long enough to witness repeated contact between those users who were regularly interacting with each other. During the period of participant-observation the researcher discovered a significant number of dead blogs. By limiting the data capture period to the most recent activity over three months, the study could focus on those users who were still active in the network. The ILB would begin to be imagined as the network of blogs that actively posted in the Irish language, the individuals who posted comments to Irish language posts in these blogs, and the relationships between them during this period. It should be noted that one blog did not include dates on its posts. Since it was impossible to verify when these posts were published, the blog was excluded from analysis, and the blogger was treated as a commenter in instances where she posted comments on other blogs during the data capture period.

ILT: The data capture period for Twitter activity was set at one month: from February 1st to 29th, 2012. Irish language Twitter users posted messages more frequently than their blogging counterparts, and the majority of the 500 most prolific Irish language accounts listed on Indigenous Tweets had tweeted in the previous month. After this cut off period the dates of users' most recent tweets dropped off rapidly. As most active Twitter accounts are updated on multiple occasions each day – aided by the ease with which tweets can be sent from mobile devices – accounts that had no new tweets in the previous month were assumed to be either dead or temporarily inactive. The decision to limit the data to just one month was also made in part because of technological factors. The computer program used to extract tweets from Twitter was restricted to approximately 3,000 tweets per account. For the most prolific Twitter users, two or three month's activity would surpass this limit.

GA: The data capture period for the *GA* Facebook group was also set at one month: from August 1st – 31st, 2012. During this period, conversations involving over 250 users were recorded and downloaded for analysis. Considering that there were 1,620 members of the group in total, it was obvious that not all members were actively participating. Despite the small sample and short data capture period, clear differences

between regular and irregular participants emerged. Over the month several new members were signed up to the group, some of whom became involved in posting messages and responding to other users. This study is only a snapshot of a month in the life of *GA*, yet within that time an evolving cluster of users was discovered. Moreover, the conversations between the active members during this period built up a significant corpus of Irish language Facebook activity for analysis.

Identifying active Irish language users

Bloggers, commenters on blogs, Twitter users, and *GA* members were only included in the study if they had posted some content (one blog post or comment, one tweet, or one Facebook message) during the data capture period where Irish was used as a significant language of communication. This was admittedly a very low activity threshold, although it is consistent with Cunliffe & Honeycutt's (2008) study of Welsh language blogs. Irish must have been used in a more than mere tokenistic manner. This ruled out those users that used Irish solely to greet their readers, or sign off messages. It also ruled out those who used Irish solely to refer to people's names, place names, or other pronouns, such as schools, teams, festivals, music and arts groups, etc. There were other conditions relating to each individual group:

ILB: The Irish language must have been used in content posted to blogs, regardless of the language used in the blog titles, descriptions, navigational tags or user profiles. In the case of photoblogs, where the primary purpose of the blog was to share photographs, the content was deemed to be Irish language if the post titles and/or photograph descriptions were written in Irish. In the case of videoblogs or 'vlogs', where the purpose of the blog was to share videos created specifically for the blog, the spoken language in the videos was checked for use of Irish. In short, blogs were only included if it was considered that the reader would have been prevented from understanding at least some of their content without engaging with the Irish language. Please note, however, that posts where Irish and English were used bilingually to express the same meaning were also included. In these instances, it is impossible to know in which language readers chose to engage with the content. However, the language choices in readers' comments would later shed some light on the commenters' linguistic engagement.

ILT: The Twitter accounts in this study were already identified as using the Irish language by Indigenous Tweets. The researcher removed from the sample those accounts that used Irish in fewer than 2.5% of their tweets according to the data on Indigenous Tweets. These were deemed to use the language too infrequently to be considered regular Irish language users. All of the tweets from the period of data capture were downloaded from the remaining Twitter accounts and checked to make sure they used the language in a more than tokenistic way.

GA: Because the moderators of *GA* maintain a strict Irish only policy (hence the group name) messages written in English are deleted from the group's wall. All the members who posted content, messages or comments to the group's wall during the data capture period were therefore considered actively engaged in interacting through Irish. 'Liking' content on the wall was not considered as Irish language activity. Although users would presumably have read other individuals' Irish language content to 'like' it, it wasn't in itself used as evidence of interacting in the Irish language.

Extracting relational data

Each web genre in the study is characterised by a unique mix of media and situation factors that define how their users communicate and interact. The social networks in this study were constructed around observable relational data between users during the data capture periods. These were defined differently in each case.

ILB: The most common form of interaction used to connect members of the ILB social network were the comments left on blog posts. Over 3,000 Irish language comments were found on blogs in the ILB during the data capture period. These were not evenly spread across blogs, however, and some blogs attracted no comments during this time. Not all comments were addressed to the blogger, but rather referenced and responded to points made by other commenters. This was interpreted by the researcher from the thrust of what the commenters had written. In these instances, comments were used to form ties between commenters interacting with each other, and not the blog author. Moreover, some commenters responded to more than one other participant in each comment and in these instances comments were used as evidence of interaction between multiple users.

Although not a very common feature of the ILB, permalinks were nonetheless used as evidence of active interaction between blogs. In reading through the blog posts and comments made to each blog during the data capture period, the researcher made note of which users created permalinks to other users and included them as evidence of ties between those users.

Trackbacks were not popularly used in the ILB, and none were found during the data capture period.

Although blogrolls formed important navigational links between blogs, it would be wrong to assume that blog authors regularly read or interacted with the blogs in their blogrolls. Most blogrolls are infrequently updated and may include blogs that once interested the author, but are now largely unread. Blogrolls in the ILB frequently included dead or dormant blogs. Although, the presence of a blogroll link between ILB members was not, therefore, considered evidence of active interaction, it was considered to form a tie between them. This is consistent with Herring *et al's* (2007) networks of bloggers hosted by LiveJournal, which used 'friend' links to tie members within each network. Like blogrolls, 'friend' links are not evidence of active interaction between members, but rather they indicate that users share some bond.

Others: As well as the four traditional forms of links described by Marlow (2004), the researcher found three other ways of linking in the ILB. These were specific to different blog hosting services.

- Individuals with profiles on the blog hosting service Blogger, can choose to follow their favourite blogs publicly. This means that an avatar, with a link to the blogger or commenter's profile, appears on the followed blog in a list of 'followers'. Much like blogrolls, the presence of followers is no guarantee that the blog is being read, and many dead and dormant blogs still had lists of followers displayed. However, 'following' was used as evidence of ties between individuals in this study.
- On blogs hosted by Tumblr, users can choose to 'like' a blogger's post or can 'reblog' the same post so that it appears on their own blog. In each case a note is added to the post being 'liked' or 'reblogged' identifying who has carried out the action. Like comments, the name of the person who created the note contains a hypertext link back to their profile. In this way the blogger can track who has been reading and engaging with their content. As such both

'likes' and 'reblogs' were used as evidence of active interaction between members of the ILB. Only one blog in the ILB was hosted by Tumblr, and only three examples of notes were found on Irish language content during the data capture period.

- In blogs hosted by Wordpress, users can choose to recommend posts by clicking on a button entitled *Mol* (or 'Like' in the English language version of the blog interface). A record is kept at the bottom of the post with the names and avatars of the people who clicked it. This is another public way for blog authors to identify some of the readers actively engaging with their content, and was used as evidence of interaction between readers and blogs. It was not commonly used in the ILB, and only five instances were found during the data capture period.

Importantly, as new blogs and commenters were discovered in comments, permalinks, trackbacks, blogrolls, followers, likes and reblogs, they were added to the blog sample and analysed for Irish language activity during the data capture period. Thus, the network of individuals involved in the ILB was expanded in the process.

ILT: Relational data between ILT members was constructed through a different methodology suited specifically to the conventions of interacting on Twitter. Following a Twitter account wasn't used as proof of interaction. Although users must actively seek and grant permission to follow each other, and the follower automatically receives content from those sites, the act of following in itself does not imply direct engagement between users. This is evident in high profile celebrity accounts that frequently attract millions of followers with whom they have little or no contact. Around 30 users in the February sample were following 1000 or more other users. The large number of accounts followed produces so much content that it can be assumed much of it goes unread. Unlike the blogosphere, most Twitter users tweeted multilingually, and it was impossible to tell which language attracted individuals to follow each account. To include ties to all followers in the network would act to expand it beyond what could truly be identified as an *Irish language* Twittersphere. A similar approach was taken by Huberman *et al* (2009) whose study of Twitter users disregarded data on followers and followees as meaningless from an interaction point of view. Instead they chose to focus on the patterns of interactions between users via directed messages. This, it was argued, produced "a social network that matters" of

people who actually communicated with each other (Huberman et al, 2009: webpage). Their study acknowledges the existence of two different networks: a dense network of followers and followees, and a sparser, "simpler" network of "actual friends" (ibid.).

There are four types of tweets a Twitter user can send: updates, @replies, retweets and direct messages. Updates are not addressed to anyone in particular; rather, they share thoughts and musings from the author to the public at large. However, @replies, retweets and direct messages are addressed to specific users and therefore formed the basis for ties in the network. On January 20th, 2012 the creator of the Indigenous Tweets website provided data for this study that made it possible to identify which users were in direct contact with each other. The data listed the usernames of all of the Twitter users who had been found by Indigenous Tweets tweeting in Irish since it began collecting information in March 2011. Beside each username the data listed all of the usernames that had been directly referred to – addressed at the beginning of @replies, included in the body of tweets as mentions, or as the source of retweeted material in retweets – in tweets that were written in Irish. The researcher removed the usernames of accounts that were not still actively tweeting in Irish during the data capture period. The remaining list, therefore, identified the Twitter users actively tweeting in Irish in February 2012 and the other active accounts they had directly referred to in Irish language tweets over ten-and-a-half months of activity in the Twittersphere.

GA: Interaction in the *GA* Facebook group was identified in the comments people posted to messages on the group wall. As in blogging, these comments built up threads of conversation over time between participants. Not all comments were addressed to the original message, but rather referenced and responded to points made by other commenters in the thread. This was interpreted by the researcher from the thrust of what the commenters had written. In these instances, comments were used to form ties between commenters interacting with each other, and not the original message's author. Moreover, some commenters responded to more than one other participant in each comment and in these instances comments were used as evidence of interaction between multiple users.

'Liking' messages is a popular form of interaction on Facebook. However, 'likes' were not included in the relational data connecting members of the *GA* group. Their exclusion was a difficult decision to make. Users would have to read the Irish

language content to 'like' it, but no Irish language response would be made. Unlike in the blogosphere, where expressing appreciation for blog posts through 'likes', 'reblogs' and the *mol* function was done very infrequently, including 'likes' in the *GA* Facebook data would have completely changed the size and structure of the group's social network. By omitting it from the data, the resulting social network only showed connections between those members who were interacting with each other regularly through Irish language text.

Collating the relational data

Source	Destination	Comments	Permalinks	Blogrolls	Follows	Tumblr likes/reblogs	Wordpress *Moltaí*
AR	GA	1	0	1	1	0	0
AR	RAM	1	0	1	1	0	0
ASD	AOT	0	1	0	0	0	0
ASD	GFG	5	0	0	0	0	0
ASD	IG	0	0	0	0	0	1

Table 4. Section of an edgelist for members of the blogosphere.

Relational data from dyads in each group (i.e. the connections linking two individuals together) was recorded in an edgelist created in Excel. The edgelist, comprising columns of usernames and numbers, identified who was linked to whom in each network. The first column identified the username of the source of interaction; the second column listed the usernames of other group members to whom they directed their interaction; and subsequent columns recorded separately in digit form the number of incidences of different types of interaction described above (see Table 4 for an example). These were directed edgelists, with the first column listing the sources of each tie and the second column listing the destinations. This would help identify reciprocal/non-reciprocal connections across the network.

2.1.2 Analysing and visualising social structure in each network

To visualise the size and structure of each network, the edgelists were exported from Excel as .txt files and imported into the statistical package R for analysis. A social network visualisation for each group was created using the Fruchterman-Reingold layout algorithm in the igraph package in R (Fruchterman & Reingold, 1991; Csárdi & Nepusz, 2006). Fruchterman and Reingold developed this force-

directed two-dimensional layout to display nodes in an aesthetically pleasing way, pulling together well-connected nodes, and pushing nodes that are less well connected to the periphery. In this way, it acts as a clear visual comparison of network density and centrality (Herring *et al*, 2007). Various functions in the igraph package in R were applied to each network to establish which nodes were core/peripheral and to analyse patterns of reciprocity in the networks.

Core nodes are defined here as those that are most globally central in each network, according to Scott (2000). This was identified by determining the degree centrality of each node in the network. Degree centrality is based on the assumption that actors with the highest degree (i.e. the most ties to other actors in the network) are the most involved and, therefore, central (Wasserman & Faust, 1994). As Freeman (1979: 219) states, a point with a relatively high degree is somehow "in the thick of things". It is a relatively simple measure obtained by dividing the degree of each node by the maximum possible degree (Freeman, 1979). It was chosen for this study above other measures of centrality – betweenness, closeness, and eigenvector centrality – in order to focus on immediate connections formed across the network. Unlike other social networks in which information may be passed along a chain of contacts, interaction in the blogosphere, Twittersphere and Facebook groups is most commonly exchanged directly between individuals in a one-to-one or one-to-many fashion. High degree centrality does not necessarily equate to high activity in each network: some relatively inactive users might be in irregular contact with many others, while more active users might focus their interactions on smaller audiences.

2.2 Analysing social position within the networks

Where possible, the researcher gathered information about the members of each network under various categories. This information would be used to explore how different aspects of user profile, location, activity and language use effected their network position. Each network was visualised with nodes of differing colours and shapes, according to the phenomena under analysis. The categories were inspired by Cunliffe and Honeycutt (2008) and their study into Welsh language blogs. Not all bloggers, Twitter users and Facebook users provide personal information about themselves in their user profiles and many remained anonymous in some of the categories. The ease with which individuals can create alter egos online – one Irish language blogger claimed to be a wombat in their profile – and the inability to verify

personal information volunteered in profiles, suggests that any personal information that was gathered should be treated cautiously. The following categories were analysed for members of each network.

Name

ILB: Blogger names were those given in the bloggers' profiles and in their comments on their own blogs and others. These were a mixture of real names and (semi)anonymous online usernames. Bloggers also commonly added their names at the end of posts. Where bloggers had no profiles, no comments and no names attributed at the ends of posts, blogger names remained anonymous. Commenter names were as they appeared at the beginning of comments. Two comments were left under the name 'Anonymous'.

ILT: Unique usernames of Twitter users were as they appeared in messages directed to them preceded by the @ sign. These were most commonly (semi)anonymous online usernames.

GA: Usernames were as they appeared at the beginning of messages posted to the group wall. These were the names that were registered with Facebook when setting up an account, and most were real-life names. Most common Christian names and surnames in Ireland have English and Irish versions, and the majority of users in *GA* used their Irish language names.

Gender

The gender of users in all three groups was identified from their usernames, user profiles or from gendered content in posts/comments/tweets/messages where possible. Users were categorised: 'Male', 'Female', 'Mixed', or 'Don't Know' ('Mixed' refers to those blogs/Twitter accounts/Facebook profiles with multiple contributors). Many individuals remained uncategorised, in particular commenters in the ILB and the Twitter accounts of companies or organisations. Because most usernames in Facebook were derived from members' real Christian names and surnames (including in Irish language names the gender specific patronyms Ó/Mac or Ní/Nic) the gender of most users could be determined.

Geographic location

This refers to where each group member was currently located. Across all three groups, geographic location fell under five categories: 'Gaeltacht', 'Ireland, possibly in the Gaeltacht', 'Ireland, outside the Gaeltacht', 'Rest of World', 'Don't Know'. 'Ireland, possibly in the Gaeltacht' refers to those users who named Ireland or one of the Gaeltacht counties as their location without specifying whether they were located inside the Gaeltacht. It should be noted that 'Ireland' refers to the island of Ireland – both the Republic of Ireland and Northern Ireland.

Across all three groups geographic location was identified from users' profiles and post/comment/tweet/message content where possible. Some users gave very specific descriptions of where they were located; others just named their country or region. Many more remained uncategorised.

User activity

ILB: For bloggers, this refers to the number of posts and comments made during the data capture period in which the Irish language was used in a more than tokenistic way. For commenters, it refers to the number of comments made to these Irish language posts regardless of language. It should be noted that where bloggers or commenters referred or responded to more than one individual in each comment they were counted as more than one comment.

ILT: The number of Irish language tweets made by each Twitter user comes from the Indigenous Tweets data, which began collecting information on Irish language users in March 2011. These figures show how many Irish language tweets were posted by each user during more than eleven months of Twitter activity. Additional information on the total number of tweets (i.e. in all languages) was recorded separately from Indigenous Tweets.

GA: This refers to the number of messages and comments made by *GA* members during the one-month data capture period. Where members referred or responded to more than one individual in each comment they were counted as more than one comment.

Language use

ILB: This refers to the predominance of Irish or English in the bloggers' and commenters' posts and comments. The categorisation was adapted from Helen Kelly-

Holmes' (2006) study into Irish language websites. It was applied to the language used in posts – in writing, audio and video – and comments, excluding direct quotes from external sources made in other languages or videos/audio embedded from other websites. The decision to exclude quotes/videos/audio from other sources was made to focus on the language of the bloggers/commenters themselves. The categories were as follows:

- IR – monolingually Irish
- Ie – mainly Irish with minor use of English words/posts
- BI – bilingual, roughly half of the content written in English and Irish
- EI – mainly English with some Irish words/posts
- Ei – English with minimal use of Irish words/posts

It should be noted the categories 'Ie', 'BI', and 'EI' differed slightly from Kelly-Holmes' study. Kelly-Holmes' 'EN' category – English only, did not apply to bloggers in the ILB, but was used for a small number of commenters who used English to comment on Irish language posts. An extra category 'SG' – Scots Gaelic – was also necessary for commenters.

ILT: Levels of Irish language use were identified from the data on Indigenous Tweets. This showed the percentage of tweets written in Irish since the website began collecting data on Irish language Twitter accounts in March 2011, and represented over eleven months of Twitter activity. Table 5 shows how the researcher equated Kelly-Holmes' categories of language use with the percentages from Indigenous Tweets. This would help create a like-for-like comparison between language use patterns in the blogosphere and Twittersphere.

Irish language use on Twitter	Helen Kelly-Holmes' categories
95-100%	IR
60-94.9%	Ie
40-59.9%	BI
20-39.9%	EI
2.5*-19.9%	Ei

*Users with fewer than 2.5% of tweets in Irish were already removed from the sample

Table 5. Kelly-Holmes' language categories in Twitter.

GA: Because the *GA* group administrators maintained a strict Irish only policy, all of the members communicated 100% of the time in Irish.

Longevity

ILB: This refers to how long each user had been active in the blogosphere. The start date for each blog was found by searching back through the blog archive and recording the date of the first Irish language post. As the commenters' first comments could not be identified, this category relied on information, where available, in the commenters' profiles on when they were first registered. Most commenters remained uncategorised, however.

ILT: This refers to when each Twitter account was first created. Each Twitter username was inputted into the online tool *When Did You Join Twitter?* (http://www.whendidyoujointwitter.com), which then provided the date and year of when that account was created. It was not possible to verify specifically when each Twitter account began tweeting in Irish, as there was a limit to how far back one could search through a Twitter feed. An assumption is made, therefore, that each user began tweeting in Irish on creating their account.

GA: Some limited information about when each group member joined *GA* could be found under the group's 'About' tab. This identified each member of the group and whether they had been members for over a year or not. For those who had been members for less than a year, it showed the length of their membership in months.

Blog/Twitter topics

ILB: This refers to the main topics of Irish language content posted to the blogs during the data capture period, as interpreted by the researcher. These were roughly recorded according to Cunliffe and Honeycutt's (2008) categorisation of Welsh blogs. Some of the more common topics included: politics, current affairs, arts and literature, and the Irish language. Both primary and secondary topics were recorded for most blogs. This categorisation did not apply to commenters.

ILT: The majority of Twitter users tweeted about their everyday lives and a mixture of personal interests. Where Twitter accounts focused on particular topics, however, these were recorded. Popular topics included 'Politics', 'Media' and 'Sport'.

GA: Topic categorisation did not apply to the *GA* group. Rather than creating their own blog or Twitter account where users could focus on their particular interests, the *GA* group maintained a public space where individuals contributed to various

thoughts, debates, event announcements, proverbs, jokes, etc. In this context, specific topics of interest could not be ascribed to individual members.

2.3 Robustness

A combination of factors was analysed to assess the robustness of each of the networks. These were informed by Herring *et al*'s (2007) assessment of robustness in language networks in the blog hosting service LiveJournal, and included the size of the networks, the presence of well-defined dense cores, and the persistence of language use across the network. In these terms it was a relative rather than absolute measure. The assessment was also made on the involvement of new members in the core of the networks, based on the hypothesis that groups that attracted new members who subsequently became regular and active participants were more likely to endure in the future.

3 Linguistic analysis of each network

If we consider the potential of the three networks in this study to be deemed online communities, we need to ask of them the same question that David Crystal asks of the web in general: "What are the shared features of language that give the […] community of users their sense of identity?" (Crystal, 2006: 6). In studying how the Irish language has been adapted to the specific websites/genres of blogging, Twitter and Facebook, corpora of discourse between core members of each network were analysed. This analysis was conducted with reference to the features explored in Cherny's (1999) register of an MUD community.

Identifying core users

Core users in each network were identified as those with the highest degree centrality. In interacting directly with multiple others across the networks it was hoped that nodes with high degree centrality would be influenced strongly by other members of the network in how they communicated. It was assumed, for example, that if language norms existed in the ILB, ILT and *GA* group they would be most apparent in the activity of these core members.

Any node in the ILB with a degree centrality of 0.19 or higher (i.e. they were connected to at least one in five network members according to the criteria of this study) was considered a core user. This resulted in a core group of 14 bloggers (and

no commenters) being set aside for further analysis. The 14 members of the ILT and *GA* group with the highest degree centralities were also identified, and their messages collated for further analysis.

Creating corpora of Irish language CMC

ILB: All of the posts and comments made in Irish during the data capture period were copied and pasted into a Word document for further analysis. This comprised a corpus of 350,000 words from the ILB over three months of activity.

ILT: A corpus of tweets posted during the data capture period by core members only was extracted using a script written in Leicester University by Dr Philip Shaw for the Twitter API (Application Programming Interface).[40] It was decided to focus on core members only because of the time-consuming nature of both extracting tweets and identifying Irish language from non-Irish language tweets in the corpus through hand-coding. Several hundred of the most recent tweets were extracted from each account and automatically collated into three files corresponding to updates/mentions, @replies and retweets (the exact number of messages extracted from each account depended on how prolifically they tweeted). This process resulted in a corpus of tweets written in Irish, English and other languages from each member of the sample. Further linguistic coding of these tweets by the researcher, resulted in a corpus of over 2,500 Irish language tweets from one month's activity and containing approximately 35,000 words.

GA: All of the messages posted by members of the *GA* group to the group wall during the data capture period were copied and pasted into a Word document. All of these messages were written in Irish. This resulted in a corpus of over 100,000 words (including the usernames and date and time stamps that had been added automatically to each message) from one month's activity.

Irish language blogosphere	Irish language Twittersphere	*GA* Facebook group
350,000 words	35,000 words	100,000 words

Table 6. Irish language corpus from each group.

[40] The researcher would like to thank Dr Philip Shaw and Dr Ruth Page of Leicester University for making the script available to this study.

Conducting linguistic analysis

The corpus of core members from the ILB, ILT and *GA* group were hand-coded by the researcher according to the linguistic features of Cherny's (1999) register of an MUD community. Linguistic features were grouped together under the following themes:

- Syntactic and morphological variation
- Abbreviations
- Play with Modality
- Language Routines

An additional theme of 'Code switching' was included in the coding process to describe the ways in which Irish language web users in the three groups switched between Irish and other languages, primarily English. The most common features from the above five themes were described and examples given. This would give a general description of the type of discourse taking place at the core of each network and the degree to which members conformed to standard written forms of the Irish language.

Dialect analysis

A separate analysis of dialect variation in the blogosphere was carried out. It was decided to focus this analysis on the ILB because of the comparatively large corpus of Irish language content extracted from the group. The objective was to explore the role of dialect in encouraging or discouraging bloggers and commenters from interacting with one another. A dialect is generally understood to refer to a linguistic variety which is grammatically, lexically and phonologically different from others (Coupland, 1988). It is most often determined on a geographical basis (Wagner, 1958). Broadly speaking, the Irish language can be divided into three main dialects today corresponding roughly to three of the four provinces of Ireland: Munster in the south, Connacht in the west, and Ulster in the north. Each dialect gives identity to the Gaeltacht regions in which it is spoken, bonding them internally and defining them externally. Such is the importance of dialect identity that the national Irish language radio station, *Raidió na Gaeltachta,* has separate studios, radio slots and news bulletins for the three main dialect regions.

English	Munster Irish	Connacht Irish	Ulster Irish
also	leis	freisin, chomh maith	**fosta**
complaint	gearán	**casaoid**, éagcaoint	éagcaoint
dance	**rince**	damhsa	damhsa
did	**dhein**	rinne	rinne
dream	**taidhreamh**	brionglóid	brionglóid
every	gach	**chuile**	**achan**
everyone	**gach éinne**	**chuile dhuine**	**ach aon nduine/achan duine**
fear	eagla	**faitíos**	eagla
for	**i gcóir**, chun	**le haghaidh**	**fá choinne**
fox	madra rua	**sionnach**	madadh rua
funeral	sochraid	sochraid	**tórramh**
harm	**díobháil**	dochar	dochar
hearing	**a chlos**	**a chluinstin**	a chloisteáil
house	**tigh**	teach	teach
How are you?	Conas atá tú? **Conas atánn tú? Conas ataoi?**	**Cén chaoi bhfuil tú?**	**Cad é mar atá tú?**
Irish (language)	**Gaolainn**	Gaeilge	Gaeilge
looking	féachaint	**breathnú**	**amharc**
minute	**nóimint**	nóiméad	**bómaite**
more	**a thuilleamh**	a thuilleadh	a thuilleadh
potatoes	prátaí	**fataí**	prátaí
quite	**cuíosach**	**réasúnta**	**measartha**
returning	filleadh	filleadh	**pilleadh**
see	feic	feic	**tchí**
seeing	**feiscint**	feiceáil	feiceáil
shout	glaoch	glaoch	**scairt**
sick	**breoite**, tinn	tinn	tinn
some	éigin	**eicint, eicínt**	éigin
table	bórd	bórd	**tábla**
too much	iomarca, iomad	iomarca	**barraíocht**
waiting	**feitheamh**	fanacht	fanacht
What is?	Cad/céard é?	Cad/céard é?	**Goidé?**

Table 7. Dialect coding scheme adapted from Wagner (1958).

Most studies into the differences between Irish language dialects focus on phonetics (see O'Rahilly, 1972; Hickey, 2011). However, since discourse in the ILB is text-based, this study focused instead on distinct lexical, grammatical or spelling features peculiar to the three main dialect regions. Here, it was inspired by the seminal dialect study of Heinrich Wagner (1958) who, in the 1950s, carried out a

survey of hundreds of terms in 88 localities around the island of Ireland, but concentrated in the Gaeltacht. His 'linguistic atlas' identified the choice of Irish words and their pronunciation in each instance, and in many cases showed clear regional differences in lexicon between the north, west and south of the island. This study focused on some key terms from Wagner's study, chosen for their likelihood to appear in contemporary Irish language discourse, and which showed clear regional differentiation. The terms peculiar to each dialect and comprising this study's coding scheme are listed in Table 7 in bold type, alongside their equivalent in the other main dialects. Some spellings have been adjusted from Wagner's 1958 study to correspond with contemporary spelling practices.

The coding scheme comprised over 40 different terms, which were searched for across the corpus of ILB discourse data. Where appropriate, the most common morphological variants of each term were included in the search. As well as the above lexical items, the researcher searched for four grammatical variations from Wagner's study: for Munster, the past tense first person singular ending *–íos*; for Connacht, the plural ending *–achaí*, and the eclipsis of certain words in western dialects after the word *sa*, or "in the" (also referenced in Ó Siadhail, 1989); and the Ulster Irish construction "le h-" used instead of the more standard "le n-" in certain circumstances in front of verbs beginning with vowels (adapted from Wagner's example *le h-ithe*, meaning "to eat"). It was hoped that these terms would help broaden the search. The use of all coding items was counted and collated for each ILB member, resulting in a general description of their dialect tendencies, if any.

English	Southern Irish	Western Irish	Northern Irish
did	dhein*	rinne	rinn
every	gach	chuile*	achan*
house	tigh*	teach	teach
How are you?	Conas atá tú?	Cén chaoi bhfuil tú?*	Goidé mar atá tú?*
looking	feiscint*	féachaint	féachaint
plural	…	-achaí*	…
potatoes	prátaí	fataí*	préataí
rather	cuíosach*/ cuibheasach	saghas/sách	rud beag
see	cíonn	feiceann	tchíonn*
table	bord/clár	bord	tábla*

*Terms already included in the 'Wagner' coding scheme.

Table 8. Dialect coding terms confirmed by Hickey (2011).

Some of the terms in Wagner's list were supported by Hickey's (2011) *The Dialects of Irish*, in which he established lexical differences between the main dialect regions of South, West and North (2011: 296 onwards). Here, South is made to correspond with Munster, West with Connacht, and North with Ulster. Table 8 lists the terms that were shared by both studies.

Critique of the dialect analysis

The dialect analysis described above is problematic on a number of grounds. Firstly, it is a rather clumsy way of drawing lines between the three main dialects; dialects that cannot be put in such watertight categories, but traditionally merged imperceptibly into each other (Ó Siadhail, 1989; Ó Cuív, 1980). The Irish of north Mayo in particular, though geographically located in Connacht, displayed many of the nuances of the Ulster dialect in Wagner's study and made identifying codes unique to Connacht difficult. Moreover, many of the features that characterised the Irish of Co. Waterford in Wagner and Hickey's studies differed from other Munster speakers.

Secondly, it is unknown how contemporary Gaeltacht residents would react to this list. Some of the terms peculiar to each region may have fallen out of popularity since Wagner's study. In many instances, Hickey refers to regional 'tendencies' towards certain features and provides examples of contemporary changes in traditional lexical features inside the Gaeltacht. Moreover, given that the official standard taught at schools contains an amalgam of regional features (Hickey, 2011: 392), supplemented by the teachers' own regional preferences, if any, it is unsure how speakers living outside the Gaeltacht, which would appear to make up the majority of ILB members, would be characterised in the analysis.

Thirdly, the analysis focusses on lexical items rather than more general phonological grammatical patterns, and as such risk stereotyping the different dialects (Wolfram, 1997). In choosing linguistic variables for analysis, Labov favours phonological features over grammatical or lexical features, as they occur more frequently and can be charted in unstructured contexts and brief interviews. There was a wide range of activity across the ILB, with some bloggers and most commenters contributing infrequently. Although the code list comprised everyday terms, it was likely that many ILB members' output would feature none of the coding items.

Finally, though some words might occur naturally in the speech of each dialect, it is by no means certain that this transfers directly to the less spontaneous context of written language in an online environment.

Despite these criticisms this coding system has some merits. It was based on Wagner's dialect study, which has stood up to academic rigour for over five decades. Moreover, some of the terms particular to each dialect were confirmed in Hickey's study of Irish from recordings with over 200 people living in the Gaeltacht today. Combining Wagner and Hickey's terms with a corpus of naturally-occurring discourse from the web at least prevented any charges of prompting the participants. Furthermore, in the small number of cases of ILB members living in the Gaeltacht, the coding results could be cross-checked with participants' geographic location to gauge its success. This dialect code list should be viewed in the context of the scale and constraints of the study, and as a first step in coding dialect in Irish language CMC.

4 Identifying online communities

Once linguistic analysis of the corpora was completed, an analysis of the three groups was conducted with reference to Herring's (2004) features of online communities. This was carried out using both social network analysis and computer-mediated discourse analysis methods.

Social network analysis

As described above (in section 2.1.2), social network analysis was used to reveal the size and social structure of the three networks, their levels of reciprocity and robustness. Further analysis focussed on the position of nodes in the network according to their geographic location, language use, subject matter, or length of activity. The results were analysed further to assess whether the groups exhibited some of the features of online communities according to Herring (2004):

- Participation structure: SNA revealed whether the networks satisfied Herring's criteria for participation structure, namely that they had a core of regular participants, actively engaged with each other through self-sustaining participation.

- Roles and hierarchy: This was partly tested through SNA, by searching for the presence of authorities and hubs within the network that could point to specific group roles. Any nodes identified as being authorities or hubs in their networks were interrogated to establish why they formed so many inward and outward ties respectively.
- Solidarity, support and reciprocity: This was partly demonstrated through SNA. As each social network was constructed around directed ties, a comparison of members' in- and out-degrees was made across the networks, which established levels of reciprocity between members.
- Shared history and culture: Shared history was established by searching through the archives of each group and then plotting the start dates of members in social network visualisations. This identified whether there existed clusters of long-established users at the core of the networks.

Computer-mediated discourse analysis

For the rest of Herring's features of online communities, discourse analysis was carried out on the corpora of core member discourse from each network. Specifically, it searched for evidence of:

- Solidarity, support and reciprocity: This was evidenced in the types of exchanges that group members shared with each other through their interactions. CMDA was used to identify examples of solidarity and support through the topics discussed in each group, the ways in which group members referred to each other, and language routines built around expressing emotions, such as encouragement, amusement and sympathy.
- Criticism and conflict: Exchanges that involved criticism and conflict were analysed to gauge whether there was a group-wide strategy for conflict resolution. This could be a formalised strategy for dealing with such incidents, or an informal one in which group members tried to diffuse the conflict in their own individual ways.
- Self-awareness: Self-awareness among group members that their group was distinct from others was established by analysing two features: addressivity and external group activities. Addressivity refers to the specific ways in which group members address the group at large. It shed light on the levels of

awareness among group members of being part of a distinct group with a perceived wider audience for their public discourse. External group activities refer to the online and offline events that group members organised to engage the group as a whole: social gatherings, seminars/webinars, festivals, etc. CMDA would also reveal examples of the 'us versus them' language proposed by Herring (2004) as one feature of group self-awareness, as well as any metaphorical sense of shared space between group members as described by Baym (2010).

- Shared history and culture: Shared culture was established through the process of linguistic analysis (described in section 3 above), which confirmed the presence of unique registers associated with each group/genre. CMDA would reveal whether the groups had explicit netiquette statements outlining expected participant behaviour (one of the criteria in Herring, 2004), and whether participants exhibited shared innovative language use (one of the criteria in Baym, 2010).

The above methodology was carried out and results were collated. The rest of this study presents these results to the reader and discusses them in the context of the themes identified earlier from reviewing literature in the fields of sociolinguistics and computer-mediated communication. Firstly, we explore the vitality of the Irish language online compared to other European minority languages.

V THE IRISH LANGUAGE ONLINE

A. IRISH LANGUAGE USE ONLINE COMPARED TO OTHER MINORITY LANGUAGES

This study compared various measures of online activity in Irish and four other minority languages in Europe. These measures shed light on the response of each language community to various web genres, and aided the researcher in assessing the languages according to UNESCO's grades of response to new domains and media, from dynamic to inactive. To put online activity in the context of real world speaker populations, it began with a comparison of user demographics in the real world.

Language	Population (Ethnologue)	Population (UNESCO)	Population (Wikipedia)[41]	UNESCO degree of endangerment
Welsh	537,870 (1991)	750,000[42] (2001)	770,000[43] (2004)	Vulnerable
Basque	658,960 (1991)	660,000[44] (1991)	665,000[45] (2006)	Vulnerable
West Frisian	467,000 (2001)	350,000[46] (1985)	467,000[47] (2001)	Vulnerable
Irish	391,470 (1983[48])	44,000[49] (2007)	2,000,000+[50] (2011)	Definitely endangered
Scots Gaelic	66,780 (2003)	58,552[51] (1993)	95,000[52] (2001)	Definitely endangered

Table 9. Minority language populations across a number of sources.

[41] Data from the English language Wikipedia pages describing each language in July 2012. These figures have been uploaded by contributors to the site citing a variety of different sources.
[42] Arwyn Watkins, T. (1993), 'Welsh' in Ball, M. J. & Fife, J. (Eds.) *The Celtic* Languages. London: Routledge.
[43] This figure is compiled using data from a number of different studies: '2004 Welsh Language Use Survey: the report', UNHCR, retrieved 5/6/2012; 'Refworld | World Directory of Minorities and Indigenous Peoples - United Kingdom: Welsh', UNHCR, retrieved 23/5/2010; 'Wales and Argentina', Welsh Assembly Government (2008), retrieved 23/1/2012 from *Wales.com website*; 'Table 1. Detailed Languages Spoken at Home and Ability to Speak English for the Population 5 Years and Over for the United States: 2006-2008 Release Date: April, 2010', United States Census Bureau, retrieved 2/1/2011; '2006 Census of Canada: Topic based tabulations: Various Languages Spoken (147), Age Groups (17A) and Sex (3) for the Population of Canada, Provinces, Territories, Census Metropolitan Areas and Census Agglomerations, 2006 Census - 20% Sample Data', Statistics Canada, retrieved 3/1/2011.
[44] R. L. Trask: The history of Basque. London: Routledge, 1997. | José Ignacio Hualde & Jon Ortiz de Urbina: A grammar of Basque. Mouton Grammar Library 26; Berlin: Mouton de Gruyter, 2003.
[45] This figure is from data cited on Wikipedia.org from the following source: *IV. Inkesta Soziolinguistikoa* Gobierno Vasco, Servicio Central de Publicaciones del Gobierno Vasco 2008, ISBN 978-84-457-2775-1
[46] Pieter Meijes Tiersma, 1985. Frisian Reference Grammar. Foris Publications, Dordrecht-Holland/Cinnaminson-USA. Streektaal.net http://taal.phileon.nl/index.php
[47] This figure comes from Dutch census in 2001.
[48] This figure appears to come from a combination of 1981 census results and Ó Riagáin & Ó Gliasáin's (1984) national language survey.
[49] Mac Eoin, G. (1993), 'Irish' in Ball, M. J. & Fife, J. (Eds.) *The Celtic* Languages. London: Routledge.
[50] This figure is from data cited on Wikipedia.org, that includes 1.77 million speakers who claimed to be Irish speakers in the 2011 census (http://www.cso.ie).
[51] Gillies, W. (1993), 'Scottish Gaelic' in Ball, M. J. & Fife, J. (Eds.) *The Celtic* Languages. London: Routledge.
[52] This figure is from data cited on Wikipedia.org including an upper estimate of speakers with some competency in Scotland and Canada.

Demographics

Comparing the speaker populations of different languages is highly problematic for a variety of reasons. Figures are usually derived from censuses from different years, some of which ask individuals simply what languages they speak; others ask them to specify how frequently they speak these languages; some ask individual questions relating to competencies in speaking, reading and writing; while others might use a combination of these questions. The value of the resulting self-assessment in painting a picture of real language use is at risk of individuals under/over-stating their language use and skills. With that caveat, Table 9 shows a number of estimates of speaker population for the five minority languages in this study, taken from three different sources. The first set of figures comes from Ethnologue,[53] a publication that lists statistics on thousands of languages from around the world, and is referenced in many linguistic studies. The Ethnologue data is not from up-to-date sources: Table 9 lists the year of data collection from the country with the largest speaker population in brackets beneath each figure. As such these figures are not an accurate representation of the numbers of people speaking these languages today. However, they do offer an estimate of worldwide speaker numbers for each language, compiled by one authoritative source, which can be used as a comparison of relative demographic strength.

The other two sets of figures in Table 9 come from the UNESCO atlas of endangered languages;[54] and from figures cited by Wikipedia,[55] which uses data posted by multiple contributors and citing a number of more up-to-date studies. While figures relating to the number of Welsh, Basque and West Frisian speakers are *relatively* consistent across at least two sources, there is a marked difference in the number of Irish speakers. This is indicative of the difficulty in defining what constitutes a language 'speaker'. Some studies provide figures for 'native' or 'fluent' speakers only; other studies draw a distinction between 'primary' and 'secondary' speakers; while others differentiate between L1 and L2 speakers. To illustrate the point, the census data from the Central Statistics Office in the Republic of Ireland[56] can be used to interpret the number of Irish speakers as any of the following: the

[53] Population according to data from Ethnologue, http://www.ethnologue.com. Some of this data comes from outdated sources from as far back as 1983.
[54] Populations as listed on UNESCO's atlas of endangered languages, sourced 23/07/2012 from: http://www.unesco.org/culture/languages-atlas/index.php?hl=en&page=atlasmap
[55] Population sources cited by Wikipedia on 12/07/2012.
[56] Data available at http://www.cso.ie.

number of people across the State over the age of three claiming that they 'can speak Irish'; the number of people across the State who speak Irish on a weekly basis, or on a daily basis, alternatively inside or outside of the education system; the number of people who live in the Gaeltacht; the number of Gaeltacht inhabitants who 'can speak Irish'; or the number of Gaeltacht inhabitants who speak Irish on a weekly basis, or on a daily basis, alternatively inside or outside of the education system. The resulting range of figures from the 2011 census could put the number of 'Irish speakers' in the Republic of Ireland anywhere between 1.77 million and 17,955.

The UNESCO figure for Irish is clearly at the conservative end of this spectrum, with a rather narrow definition of what constitutes an 'Irish speaker' seemingly based on the number of self-identified Irish speakers living in the Gaeltacht. The resulting estimate shows fewer Irish speakers than Scots Gaelic. This figure contrasts markedly with the one cited by Wikipedia, which gives a far more optimistic view of the number of Irish speakers worldwide at more than two million. The Wikipedia figure is equally unrealistic an estimate as it counts anyone who claims some competency in the language as a speaker, regardless of how often they use the language or whether they live in communities where the language is commonly spoken. Their figure for Scots Gaelic is also markedly higher than both Ethnologue and UNESCO.

Table 9 paints a general picture of five languages with a spectrum of demographic strength. Excluding the seemingly overly-pessimistic and overly-optimistic figures for Irish language speakers, one can generally rank the languages' demographic strength accordingly: Welsh and Basque are relatively robust with around 700,000 speakers each and are alternatively ranked first or second in the table; West Frisian and Irish are weaker with fewer speakers and come either third or fourth; and Scots Gaelic is the least spoken. This is generally reflected in UNESCO's rating of the languages according to their degree of endangerment, with Welsh, Basque and West Frisian described as 'vulnerable' and Irish and Scots Gaelic classified as 'definitely endangered'.

The classification process is an interpretative one, and the status of Irish as 'definitely endangered' seems to jar with the language's official status both at a national and EU level, its compulsory study in primary and post-primary schools, and the Irish Government's continued investment in Irish language media and services. It is fair to say, however, that the language has suffered a history of decline and peripheralisation dating back at least to the 16th century (Mac Giolla Chríost, 2005),

to the point where today its status as a community language in its traditional heartland of the Gaeltacht is under threat. The statistic from the 2011 census of just under 18,000 Gaeltacht inhabitants speaking the language on a daily basis outside the educational system is a particularly worrying one. Other studies into the language's use indicate a low rate of intergenerational transmission within Gaeltacht families that further threatens the vitality of the language in its heartland (Ó Giollagáin *et al*, 2007). Moreover, in the context of recent changes in government policy towards the Gaeltacht aimed at saving the language from further decline in these regions, the classification of 'definitely endangered' does not seem overly pessimistic. Language activists, however, might point to the sustained growth in Irish-medium education elsewhere on the island of Ireland and the rise in the numbers of Irish speakers in towns and cities outside the Gaeltacht to counter a more positive assessment of the language's vitality.

This study tested UNESCO's ratings of each language's degree of endangerment by comparing online activity in each language. This process involved reconsidering the various figures on demographic strength above in comparison to how actively the languages were used online. To this end the study now turns to measures of *user profile*, *web presence* and *user activity* to build a more comprehensive picture of language use online.

User profile

This approach assesses the number of speakers of each language that has access to the internet. We have already discussed the demographic position of each language and the difficulties in identifying who is a 'speaker' from census data. For the sake of comparison, the estimated numbers of speakers listed in Table 10 have been taken from the least extreme or (in the case of two sources citing the same figure) the most common population estimates from the three sources in Table 9. In the case of Irish, this favours Ethnologue's figure of approximately 390,000 speakers worldwide: a figure which is largely derived from a national language survey conducted by *Institiúid Teangeolaíochta Éireann* in that year and estimating numbers of 'fluent speakers' in the Republic of Ireland (see Ó Riagáin & Ó Gliasáin, 1984). The figures in Table 10 are presented alongside the percentage of households with access to the internet in the countries with the largest speaker populations. Even on a national level these figures – sourced from the Europe-wide statistics agency Eurostat – are

problematic, as they do not include online access via mobile phones or other devices other than the home computer. Moreover, the data compares the percentage of households, rather than the percentage of the population who have internet access. The Eurostat data indicates that the Netherlands and the UK are among the best-connected countries in Europe when it comes to internet access; while Spain lags behind its European counterparts. An estimate of the user profiles of each language is derived by multiplying the estimated speaker population by the percentage of households with internet connections. According to this crude figure, West Frisian has the second highest number of potential web users, more than Basque. However, as we'll see, the potential user profiles in Table 10 do not correspond to actual user activity in the featured languages on some of the web's most popular sites.

Language	Estimated number of speakers[57]	Country with largest speaker population	Internet Penetration [58]	Potential online population
Welsh	750,000	United Kingdom	85%	~ 637,500
Basque	660,000	Spain	64%	~ 422,400
West Frisian	467,000	Netherlands	94%	~ 438,980
Irish	391,470	Republic of Ireland	78%	~ 305,350
Scots Gaelic	66,780	United Kingdom	85%	~ 56,760

Table 10. Internet penetration and potential number of speakers online.

It should be noted that these languages, in particular Scots Gaelic, are traditionally spoken in some of the most peripheral regions of their native countries and the internet penetration in these regions is likely to be below the national average. Details of internet penetration on a regional level are often difficult to find. Ireland's Central Statistics Office gathers lots of data on internet usage, which is then made public on their website. This includes specific data on the Gaeltacht, as it is designated statutorily. For example, using their Small Area Population Statistics[59] to isolate census data from the Gaeltacht areas, it was found that on average 66.6% of Gaeltacht homes had internet access compared to 73.6% nationally. The gap was even more pronounced in the more remote rural areas where Irish is strongest: like in Carna in

[57] Devised by taking the middle/most common estimate from the three sources in Table 1.
[58] Internet penetration data from Eurostat. See Seybert, 2011 in bibliography. Data refers to internet penetration at a national level according to the main country of residence of language users.
[59] Data sourced from the Central Statistics Office website on 01/08/2012 from: http://census.cso.ie/sapmap/

Co. Galway where 57.7% of households had internet access; or Gort an Choirce in the Co. Donegal Gaeltacht where the figure was just 53.7%. As well as occupying some of the most peripheral regions of Ireland, the Gaeltacht is home to a higher than average number of older people, who in general are less likely to use the internet (CSO, 2011: 11). According to the 2011 census, over-65 year olds made up 14.8% of the Gaeltacht population, compared to just 11.7% on a national level. The spread of broadband and wireless internet services in rural areas in particular comes at a slower rate than their urban counterparts due to the reduced commercial viability of extending these services to remote areas with less demand. However, projects such as the recent investment by telecommunications company Eircom in wireless access in Dingle town (in the County Kerry Gaeltacht) shows how collaboration between state and commercial forces can improve access to web services in peripheral areas.

Language	Wikipedia	Facebook	Twitter	Blogger.com	Google Translate
Welsh	Yes	Yes	No*	No	Yes
Basque	Yes	Yes	No*	No	Yes
Irish	Yes	Yes	No*	No	Yes
West Frisian	Yes	Yes	No	No	No
Scots Gaelic	Yes	No	No	No	No

*In the process of being translated.

Table 11. Availability of major web services/interfaces in Sept 2012.

Web presence

Web presence refers to the number of websites available in each language. As Table 11 shows, there were differences in the availability of some of the most popular web sites and services in the five minority languages. Wikipedia is a crowd-sourced online encyclopaedia where users from across the world freely upload and amend articles on a vast range of topics. There are editions of Wikipedia in all five of the featured languages contributed to and updated by speakers of those languages, albeit with widely varying numbers of pages and amounts of content.

Welsh, Basque, West Frisian and Irish language versions of Facebook are also available. The translations used in these interfaces are crowd-sourced from users of the site who contribute and vote on the most appropriate versions of common Facebook phrases. Scots Gaelic was the only language from this study that was not available as a Facebook interface. Facebook users are free to communicate in whatever languages they choose, as long as the site supports the languages' alphabet

systems. Thus, despite the lack of a Scots Gaelic interface, there are examples of Scots Gaelic groups where members interact publicly in the language; and Facebook users can choose to interact with each other in Scots Gaelic on an individual basis on their friends' walls or in private messages, albeit using an English (or other) language interface.

Twitter also crowd-sources translations from its users, although in September 2012 this was limited to just 30 of the most widely used languages, including Catalan. As of September 2012, none of the languages under study were available in the Twitter user interface. However, during that month the Twitter Translation Centre was opened to sixteen new languages including Basque, Irish and Welsh,[60] and Twitter interfaces for these languages are in the process of being translated by users.

None of the five minority languages were featured among the fifty interface languages for the popular blog hosting service Blogger.com, despite the service being used to publish many blogs in these languages.

Significantly, Welsh, Basque and Irish are included among the 65 languages available in Google Translate. This means that texts in Welsh, Basque and Irish can be translated directly into 64 other languages and *vice versa*. Despite the frequent errors and mistranslations, particularly among lesser-used languages, this is an extremely valuable language resource. West Frisian and Scots Gaelic users cannot avail of this service in their language, reinforcing the languages' peripheral status online.

Table 11 already points to a spectrum of availability of popular online services in the five languages. Although it doesn't indicate how many people use the sites in each language, or how much content is produced on them, the availability of templates does point to a perceived level of demand among some of the biggest online brands worldwide for services in these languages. Basque, Irish and Welsh appear to share a common profile of availability, while West Frisian and Scots Gaelic are less well represented.

Table 12 shows the results of searches on three of the world's most popular web search engines: Google, Yahoo! and Bing. The search terms "Olympic Games", "European Union" and "Solar System" were acquired from the titles of Wikipedia pages on the topics in the five languages, and searched for without using inverted

[60] See blog entry 'Language on the Go' at http://translate.twttr.com/cms/IntlBlog.

commas, Boolean commands or alternative morphological variants. The researcher selected these terms on the basis that they would be recognised globally and commonly used by speakers in the languages. It was hoped that the combined results would offer some comparison of web presence. Table 12 shows the resulting numbers of hits, ranging from over 425,000 to just 112.

Language	Google search for: Olympic Games	Yahoo! search for: European Union	Bing search for: Solar System
Basque	425,000	13,800	9,630
Welsh	98,500	14,200	413
Irish	62,100	9,290	295
West Frisian	18,500	478	159
Scots Gaelic	NA[61]	236	112

Table 12. Search results for each language (July 2012).

Although the numbers of hits varied widely between search engines and languages, the results would appear to confirm – with one exception (see footnote 61) – the relative demographic strength of the five languages. Basque and Welsh came alternatively in first and second place, albeit with large differences in the numbers of hits; Irish and West Frisian had fewer hits; and Scots Gaelic returned the fewest hits for each search term. In all three searches, Irish hits outnumbered those in West Frisian. Considering its demographic strength relative to Irish according to some sources and it being rated as less endangered by UNESCO, the fewer hits in West Frisian came as a surprise. It should be noted that though the web pages quantified in Table 12 were found as a result of searching for minority language terms it does not necessarily comply that they were published in those languages. For example, Kelly-Holmes (2006) conducted five Google searches for common Irish terms and analysed the first ten hits from each search. She found that over 30% of the hits were either mostly or completely written in English.

Some studies refer to numbers of articles on the online encyclopaedia Wikipedia as a convenient measure of the relative web presence of different languages (see Deere, 2011). As of Sept 2012, there were 285 different language editions with widely varying numbers of pages: from over four million English language pages to just one

[61] The Scots Gaelic term for Olympic Games – Na h-Olympics, according to the dedicated Wikipedia page on the topic – was too close to the English word and returned 92,600,000 results, which is not an accurate representation of Scots Gaelic's web presence.

page in the Herero language native to Namibia and Botswana. Table 13 shows a large difference between the numbers of articles published on Wikipedia in the five minority languages. Basque had a considerably stronger presence with approximately 138,000 articles, almost four times as many articles as Welsh; Welsh in turn had approximately 37,000 articles; West Frisian had 25,000; and Irish and Scots Gaelic had around 16,000 and 9,000 articles respectively. The relatively large number of Basque articles indicates the website's general popularity among Basque speakers.

Language	Wikipedia articles (Sept 2012)	Wiktionary articles (Sept 2012)
Basque	138,368	35,248
Welsh	37,492	14,520
West Frisian	25,484	13,323
Irish	15,685	2,336
Scots Gaelic	8,941	174

Table 13. Minority language Wikipedia and Wiktionary use.

As each article may contain alternatively a very long and detailed analysis or a short synopsis of the topic, the number of articles alone does not indicate the amount of Wikipedia content in a given language. To take one example, entries for the "Olympic Games" in July 2012 included articles with approximately 600 words for Basque and Welsh, 650 for Irish, 850 for Frisian, and just 50 words for Scots Gaelic, compared to over 15,000 words in the English language edition.

Andrew Deere (2011) of the University of Glamorgan carried out a detailed assessment of the presence of twenty languages – including Welsh, West Frisian, Irish and Scots Gaelic – on Wikipedia. Taking a number of factors into account, including the number of articles in each language; the average number of words, images and links in each article; and adjusting for languages in which it takes more words to express the same content – he assigned a score to each language according to their relative presence on the online encyclopaedia. Although Basque was not included in his study, the scores reflect the ranking assigned to the languages in Table 13, with Welsh having the largest presence of the four languages, followed by West Frisian, then Irish and lastly Scots Gaelic.

The ranking of languages in Wikipedia was mirrored in the number of Wiktionary articles published in each language. Wiktionary is a multilingual crowd-sourced online dictionary and a sister site of Wikipedia. As of July 2012 it was available in

158 different language editions. As Table 13 shows, there is a large difference between the number of words translated in each language, from over 30,000 in Basque to just 174 in Scots Gaelic. Although the numbers confirm the relative strengths of the languages ranked demographically, it should be noted that some of the languages have already established popular online dictionary sites. This might explain the small number of Wiktionary entries in the Irish language, for example. There are already some well-established high quality Irish dictionaries online – developed primarily by Irish universities – that may be deemed to serve the language community adequately without recourse to Wiktionary. This is an example of effective institutional support for the language in an online context.

Language	Registered Wikipedia users	Active Wikipedia users*	Prolific Wikipedia users†	Page views per month
Basque	38,919	284	1,086	4,983,604
Welsh	22,322	106	393	3,226,559
Irish	15,061	83	320	1,224,360
West Frisian	12,694	85	183	1,370,287
Scots Gaelic	8,364	53	136	882,820

*Users who have edited at least once in the past 30 days
†Users who edited at least 10 times since they arrived

Table 14. Wikipedia users in September 2012.

User activity

The third method of measuring language activity online is by focusing on the users and their actual language use on individual websites. As well as the numbers of articles currently in each language edition, Wikipedia also gives details of the number of users – both registered and active – and the number of page views in each language version, as displayed in Table 14. Here the difference between the number of registered and active users across the languages is not as pronounced as the difference in the number of articles, indicating that per capita Basque users are particularly prolific when it comes to publishing articles on new topics. Basque has the highest number of registered and active users, followed by Welsh, then Irish, West Frisian and Scots Gaelic. The data shows that Wikipedia is regularly being updated in all five minority languages by varying numbers of users. The figures do not reflect the relative demographic strengths of the languages according to the sources cited by Ethnologue, UNESCO and Wikipedia, however, with Basque users far outnumbering those in Welsh. Irish users unexpectedly outnumber West Frisian, although the

number of views attracted by West Frisian pages is higher.

Language	Twitter users Mar 2011 - July 2012	Blogs Sept 2011 - July 2012
Basque	13,041	NA
Welsh	8,278	261
Irish	3,166	118
West Frisian	2,038	17
Scots Gaelic	375	56

Table 15. Minority language users in different social media according to Indigenous Tweets and Indigenous Blogs.

Table 15 shows the number of Twitter users found to have tweeted in the five languages between March 2011 and July 2012 as listed on Indigenous Tweets. Although individual social media websites are more popular in some countries than others, the number of active Twitter users in the featured languages roughly reflects the numbers of individual users uploading content in these languages to Wikipedia: with Basque having by far the most active users; followed by Welsh; with Irish and West Frisian having similar numbers of users; and Scots Gaelic having the fewest users.

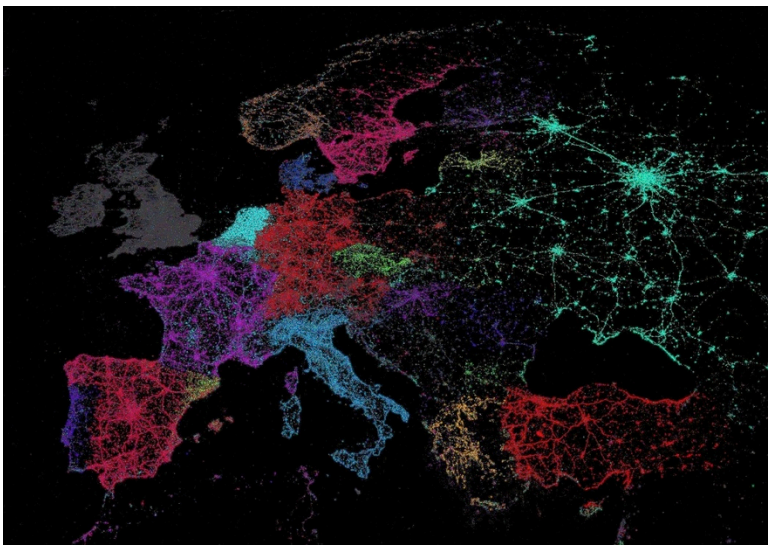


Fig. 3: Map of Europe indicating the languages tweeted from locations across the continent. © Eric Fischer

To put these numbers in a wider context, Fig. 3 shows a map of Europe indicating the languages tweeted by individuals across the continent over a five-month period in

2011. It was developed by computer programmer and map enthusiast Eric Fischer[62] using a web crawler and Google's language detection software to identify Twitter users' individual locations and languages. The result is a colourful representation of the range of languages tweeted by Europeans. Some national languages, such as Portuguese and Danish, are clearly defined by their country's borders; other supranational languages, such as French and German, cross political borders into the regions of neighbouring countries in which they are spoken. On a regional level, Catalan is clearly visible in its traditional heartland of northeast Spain, a fact picked up on by the Catalan Government and used to promote the language's vitality in the region.[63] However, none of the minority languages in this study showed up on the map. Welsh, Irish and Scots Gaelic, which were included in the study,[64] remain hidden under the dominance of English even in their traditional heartlands. Neither Basque nor West Frisian is visible on the map either.

Table 15 also includes the numbers of blogs in Welsh, Irish, West Frisian and Scots Gaelic as listed on Indigenous Blogs. Basque is not included on this site, and the blogs listed are those hosted by Blogspot (Blogger.com) only. As the site's creator Kevin Scannell reasons in a post to the blog accompanying the site:[65] "[Blogspot] hosts more than 90% of the blogs written in languages I'm interested in". In general, the number of blogs appears to reflect the languages' relative popularity in other social media with Welsh having the most blogs, more than twice the number of Irish blogs, which has in turn more than twice the number of Scots Gaelic blogs. West Frisian is significantly under-represented on Indigenous Blogs and this might be because of alternative blog hosting services being more popular in the Netherlands.

Response to new domains and media

The figures shown above were selected from a few popular sources and are a convenient and overly simplistic assessment of online activity in general. However, they do offer some insight into the relative activity in each language online. Users of Welsh, Basque, West Frisian, Irish and Scots Gaelic will undoubtedly continue to

[62] Image from Eric Fischer's Flickr account, sourced 22/03/2012 from: http://www.flickr.com/photos/walkingsf/6277163176/in/photostream.
[63] See 'Un mapa reflecteix la vitalitat del català a Twitter', sourced 11/04/2012: http://www20.gencat.cat/portal/site/Llengcat/menuitem.21576464db9e81e7a129d410b0c0e1a0/?vgnextoid=035748a344371110VgnVCM1000000b0c1e0aRCRD&vgnextchannel=035748a344371110VgnVCM1000000b0c1e0aRCRD&vgnextfmt=detall&contentid=e173d8d408883310VgnVCM2000009b0c1e0aRCRD
[64] This was confirmed in a personal communication with the map's creator.
[65] The post can be viewed here: http://indigenoustweets.blogspot.ie/2011/09/new-feature-indigenous-blogs.html

adapt their languages to computer-mediated communication online. All five languages have a presence in the websites under study here – Wikipedia, Wiktionary, Twitter and blogs – although the interfaces were not always available in all languages.

It would be impossible to grade each language's response to the web in a robust way without further detailed analysis. However, using the vocabulary of the UNESCO grading system to describe the language communities' response to these new forms of CMC, it is the opinion of this researcher that both Basque and Welsh can be described as having a robust and active presence online. Both have groups of users actively blogging, tweeting and publishing Wikipedia and Wiktionary articles at levels that are consistently higher than the other three languages. The activity of West Frisian and Irish language users certainly points to language communities that are receptive to online use, albeit with fewer active users than in Basque or Welsh. Finally, Scots Gaelic has a relatively small presence online. Although there are a small group of users publishing tweets and blogs in the language, as well as contributing content to the Scots Gaelic Wikipedia and Wiktionary editions, the overall online activity of the language might best be described as 'coping' according to UNESCO's grading system.

The rest of this study focused solely on the Irish language and its use online. Firstly, an inventory of Irish language websites was conducted. This helped identify the most popular genres for Irish language discourse online, which in turn informed which web genres would be the focus of further analysis.

B. POPULAR IRISH WEB GENRES IN IRISH

Web inventory by language use

The inventory of web pages where Irish language content might be found comprised over a thousand different URLs. Table 16 shows the results of the researcher's hand-coding of each web page for Irish language content. This coding process involved searching through the websites listed in the Irish language web portals/authorities and the top 150 popular websites in Ireland on Alexa, and scrolling through the web pages found through Google searches, to establish the level of Irish language use in each case.

Total web inventory	1,153 websites/pages
Significant Irish language use	742 websites/pages
Little Irish language use	38 websites/pages
Computer-generated Irish language	165 websites/pages
No Irish language use	202 websites/pages
Could not be coded	6 websites/pages

Table 16. Web inventory according to Irish language content.

Approximately 65% of web pages in the inventory were found to have content that pointed to more than just occasional or token Irish language use. These were coded as having 'significant' Irish language use. The most fruitful sources of Irish language content were the Irish language web portals and authorities. As expected, the vast majority of websites listed by Gaelport, Sabhail Mòr Ostaig, Ní Chartúir's (2002) book *The Irish Language: An Overview and Guide*, and the Irish Government's Department of Community, Equality and Gaeltacht Affairs website contained significant Irish language content. This suggests an important role for language organisations in promoting minority language web content, and represents another example of institutional support for the Irish language online. On the other hand, most of the top 150 popular websites in Ireland listed by Alexa – over 70% of sites that could be coded – contained little or no Irish language content. These were primarily corporate websites – both national and international – and usually operated entirely through English. This suggests a perception among business leaders in Ireland that there is little demand or business imperative for Irish language content and services online. This assumption is presumably made in the context of there being no monolingual Irish speakers in Ireland, despite the 1.77 million (or approximately 40% of people over the age of three) who claimed they could speak Irish in the 2011 census.

There were mixed results from the Google searches, with many web pages (28.4% of the sample) containing little or no Irish language content despite the search terms being in Irish. One unexpected result from these searches was the large number of pages – over 200 in total – with computer-generated Irish language content. These included a Pakistani web design company that advertised its services in Irish;[66] a current affairs blog from the Dominican Republic that published some posts in Irish;[67] and an Australian national newspaper that had a web version translated into Irish.[68]

[66] See http://www.ire-s.com/ga/tag/flash.
[67] See http://digitalgroup.info/wordpress/index.php/archives/95444.
[68] See http://blogs.smh.com.au.ga.mk.gd/entertainment/tabletalk.

There were also over a hundred discussion forums on various themes and topics that provided Google-powered automatic translation into a number of different languages.[69] The researcher found no evidence on these discussion forums of users writing original content in Irish. Rather, Irish appeared to be included in the list of computer-generated language versions as a formality. Without exception the content had been poorly translated and contained frequent grammatical errors, jumbled syntax and the inclusion of non-Irish words in instances where Irish translations did not exist or the original content was misspelt. The Meta-Net study into language technology support for the Irish language (Judge *et al*, 2012) concluded that Irish has weak or no support in machine translation. This is common among many of Europe's minority and lesser-spoken languages. However, as more investment is made in translation software, and as the corpus of Irish language texts on which the translations are made grows, machine translation might have the potential to provide more high quality Irish language content in the future.

Web inventory by source

The inventory contained examples of multiple web pages from the same websites that had been found at different stages of the web search. Once duplicates were removed from the list of 742 web pages with significant Irish language use, the remaining 496 websites were categorised as 'Official', 'Commercial', 'Third Sector', or 'Personal Interest' according to the sources of Irish language content. This was an interpretive process, and it was not always clear who had uploaded the Irish language content and for what purposes. Figure 4 shows the share of websites in each category. Each category was well represented, suggesting that the methodology used to source Irish language websites had been well constructed. However, of the three websites that were found separately by the researcher as he browsed the web, and later used to test the web searches' validity, only two were included in the inventory. The *People's Republic of Cork* website, which has an Irish language forum, was not found in the search. Thus, although the inventory can be said to include a wide spectrum of Irish language websites, it is certainly not an exhaustive list and should be considered a sample rather than a complete set.

[69] For an example, see http://www.vbenterprisetranslator.com/forum/ga/general-discussions/1410-iarratas-plé-cad-mar-gheall-ar-leagan-xenforo.html.

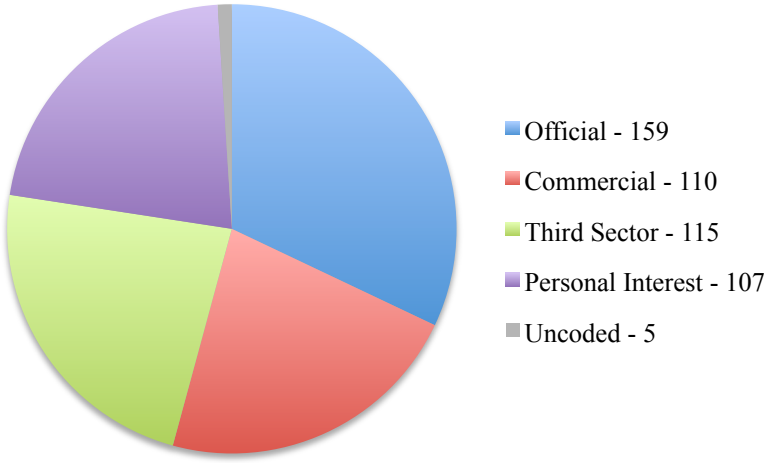


Fig. 4. Origins of Irish language websites.

The largest source of websites was official organisations, representing state agencies in Ireland (at local, national and European levels), public service broadcasters, and state-funded schools and universities. *Gaelscoileanna*[70] were particularly well represented, with 18 different schools found posting information in Irish on their websites. Universities inside and outside of Ireland were also well represented in this category.

The second largest category was third sector organisations, representing organisations that received some funding through government agencies to promote Irish language and culture. At one end of the spectrum, these included large national organisations aimed at promoting Irish-medium education; traditional Irish music, singing and dance; as well as online archives of Irish cultural artefacts. Smaller organisations included local cultural centres; arts and literature festivals; and Irish language clubs and societies. The dominance of official and third sector websites in this study reflects the findings of Kelly-Holmes' (2006) study, where official and other sectors directly affected by language policy and planning were found to produce the highest quantity of Irish language content.

The commercial sector was unexpectedly well represented with over 100 examples of businesses communicating online through Irish. Most of these offered products and services related to the language itself: language courses, translators, Irish language publishers, online bookshops, traditional Irish music and singing, and computer software tailored to the Irish language. There were 12 Irish language summer colleges

[70] The plural of *Gaelscoil*, these are state primary schools in Ireland where the curriculum is taught through the medium of Irish.

represented in the sample. Any large international companies included in the sample were online service providers like Facebook, Google and Gumtree that had Irish language users who contributed content in the language. Apart from these, the Irish Times newspaper and the mobile telecommunications provider Meteor were the only large well-known companies found using Irish on their websites. This was a disappointing result considering that the top 150 popular sites in Ireland, as listed by Alexa, were included in the search, and contained the websites of some of Ireland's biggest commercial brands.

The fourth category of 'Personal Interest' included the websites of individual language enthusiasts. These included blogs and Facebook pages in Irish, websites dedicated to Irish language groups and clubs, and an eclectic mix of personal web pages where individuals uploaded information in the Irish language on a range of topics, from ancient manuscripts to local football clubs. Most of the personal web pages were rarely updated and poorly designed, appearing old-fashioned in web design terms.

Popular Irish language web genres

The 496 websites were filtered further to identify any publicly accessible interactive websites with regularly updated Irish language content shared through a significant number of repeat users. A total of 77 websites were found to satisfy these criteria to varying degrees. These were subsequently grouped into seven primary and three secondary web genres, according to the types of discourse shared there.

Primary genres

These describe the types of websites that were found to host the most frequent Irish language interaction, and include blogs, news/magazine websites that included comment forms, discussion forums and boards, listservs, microblogs, social networking sites, and wikis. They represent the most popular sites of online discourse in the Irish language, and form the basis of the genre analysis later in the study.

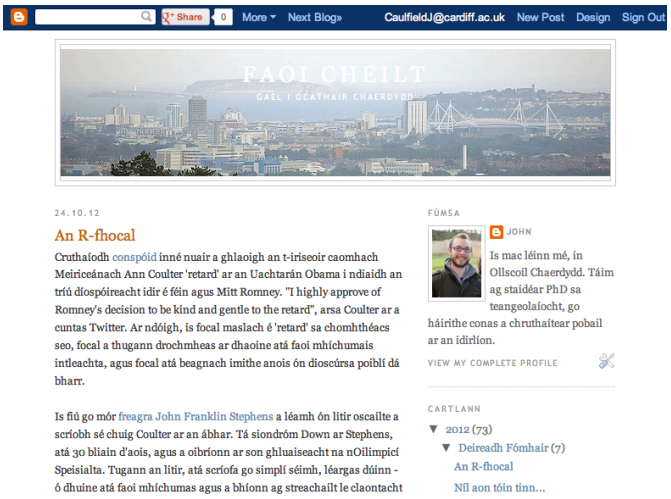


Fig. 5. Faoi Cheilt blog[71]

Blogs. The search discovered seven active Irish language blogs with varying levels of activity. These were most commonly hosted by one of two blog hosting services: Blogger and Wordpress. Only one of the blogs had frequent comments from readers. However, the dormant blogs linked to other active blogs in their blogrolls, which in turn linked to other active blogs, thereby revealing the further extent of the Irish language blogosphere.

Examples: Smaointe Fánacha Aonghusa (http://aonghus.blogspot.com)

Gasúir le Gaeilge (http://gasuirlegaeilge.wordpress.com)

Fig. 6. Beo! magazine website[72]

[71] http://www.faoicheilt.blogspot.com, screen grab sourced 26/10/2012
[72] http://www.beo.ie, screen grab sourced 26/10/2012

Comments on news/magazine websites. News and magazine websites frequently invite readers to respond to their content by providing a comment form at the end of individual posts or articles. The search found seven examples of websites where readers left comments on their articles in Irish. In the cases of primarily English language news websites, Irish language comments were very infrequent. However, in the examples of monolingual Irish news/magazine websites listed below Irish language comments were frequent and often created threads of comments over time between readers and authors.

Examples: Nós (http://www.nosmag.com)

Nuacht24 (http://www.nuacht24.com)

Fig. 7. Daltaí discussion forum[73]

Discussion forums/boards. A discussion forum or message board is an online discussion site where users contribute to conversations in the form of posted messages. The search discovered seven active discussion forums in the Irish language. Most of these websites had designated forums for the Irish language, where individuals were expected to interact in Irish only. In some cases, the Irish language forum constituted only a small part of the wider collection of English language forums/boards.

Examples: Boards.ie (http://www.boards.ie)

Second Level Support Service (http://gaeilge.slss.ie/forum)

[73] http://www.daltai.com, screen grab sourced 26/10/2012

Fig. 8. Gaeilge-A listserv[74]

Listservs. A listserv is a mailing list that allows members to send emails to all members of the group, building conversations over time on various topics. The search found evidence of four listservs using the Irish language. Two of these listservs required a log-in and did not archive their messages online. They therefore remained outside the scope of this study. However, it was possible to view the archived messages of some of the listservs hosted by listserv.heanet.ie, which showed evidence of frequent Irish language interaction.

Example: Acmhainn Email List (http://www.acmhainn.ie/foram.htm)

Fig. 9. Gaelport Twitter account[75]

[74] https://listserv.heanet.ie/cgi-bin/wa?A0=GAEILGE-A, screen grab sourced 26/10/2012
[75] http://twitter.com/Gaelport, screen grab sourced 26/10/2012

Microblogs. A microblog is a blog that contains very short posts, and it is most commonly associated with the microblog hosting service Twitter. The search found 21 examples of Irish language Twitter accounts with varying levels of activity. They in turn linked to other Irish language Twitter accounts through their lists of followers and the lists of Twitter accounts they were following.

Examples: Club Leabhar Twitter page (http://twitter.com/ClubLeabhar)

Foinse Twitter page (http://twitter.com/Foinsenuachtan)

Fig. 10. TG4 Facebook page[76]

Social networking sites. A social networking site is a site that creates social networks between users who, for example, share friendship, common interests or activities. The search found 21 examples of Irish language Facebook pages, and one Bebo page. The social networking site Bebo has since closed down. The Facebook pages were those of groups, organisations or companies that made their content open to the public. These in turn linked to other Irish language Facebook pages through their lists of friends or members. It can be assumed that the Irish language is also used by many individuals in their private networks. These remained hidden from the study, however. A specifically Irish language social networking site called AbairLeat (http://abairleat.kontain.com) was launched in February 2012, after the search for Irish language websites was conducted. As such, it remains outside this study's analysis.

Examples: Cló Iar-Chonnachta Facebook page

[76] http://www.facebook.com/TG4TV, screen grab sourced 26/10/2012

(http://www.facebook.com/cloiarchonnachta)

An Cumann Gaelach UCD Facebook page

(http://www.facebook.com/cumanngaelachucd)

Fig. 11. 'An Ghaeilge' Irish language Wikipedia 'Talk' page[77]

Wikis. A wiki is a communal website through which multiple users can access and contribute to web pages on a particular theme or topic. The most popular Irish language wiki was that of Wikipedia, an online encyclopaedia where users from around the world collaborate in contributing information on a wide range of topics and in over 250 languages. The search found evidence of Irish language pages and user accounts on Wikipedia and Wiktionary. A small number of users were found interacting in Irish on the discussion or 'talk' pages associated with some Irish language encyclopaedia or dictionary entries.

Example: Irish homepage on Wikipedia

(http://ga.wikipedia.org/wiki/Príomhleathanach)

Irish homepage on Wiktionary

(http://ga.wiktionary.org/wiki/Plé:Príomhleathanach)

Secondary genres

A number of additional genres were identified where the Irish language was used infrequently. These are described below. They are not included in later genre analysis as the level of Irish language activity on these sites was deemed too irregular.

[77] http://ga.wikipedia.org/wiki/Plé:An_Ghaeilge, screen grab sourced 26/10/2012

However, they illustrate the potential for the Irish language to be used across a range of outputs.

Google groups/Yahoo! groups. These are discussion groups hosted by Google and Yahoo!, which enable users to interact with each other around commonly shared interests. They differ from discussion forums and boards because they are composed of thousands of individual groups each with their own varied topics of interest. Also, discussion group members can choose to read and respond to messages from other individuals via email. The search did not identify any specific Irish language web groups. However, additional browsing on the Google groups and Yahoo! groups websites found a number of Irish language groups. These were not frequented by many users, however, and were updated irregularly.
Examples: Google groups (http://groups.google.com)
Yahoo! groups (groups.yahoo.com)

Video-sharing websites. Websites such as YouTube and Vimeo enable users to upload and share videos online. These users can create channels through which all of their uploaded videos can be accessed. Viewers can then interact with the uploader and other viewers by adding comments about the videos. These comments may build into threads of conversation over time. The search found four examples of channels on video-sharing websites – primarily YouTube – that included Irish language videos and interactions. These were irregularly updated and attracted infrequent comments, however.
Examples: Nuacht24 YouTube channel (http://www.youtube.com/user/nuacht24)
TG Lurgan Vimeo channel (http://vimeo.com/lurgan)

Others. The search also found instances of infrequent Irish language interaction on websites that did not fit into the genres described above. These were on the image-sharing website Flickr; the question and answer website Answers.com; an interactive map on the BBC Irish language microsite; and the location-based social networking site foursquare. None of these were being used interactively on a regular enough basis to be included in later analysis. However, they do show the potential for the Irish language to be used creatively in a wide spectrum of web sites and services.

Genre analysis

This study describes the primary web genres according to how they were observed being used through the Irish language. Medium factors may vary widely across genres, and the descriptions below pertain specifically to the Irish language sites in this study. So, for example, while some discussion forums elsewhere on the web may have specific ways of quoting other people's comments, most Irish language forums did not. Likewise, situation factors may vary across genres, especially in the topic and tone of content. So, for example, whereas some listservs elsewhere on the web may exhibit more informal interactions, the Irish language listservs were observed to exhibit, on the whole, more formal interactions involving knowledge exchange, albeit peppered with humour and casual interactions when the topic allowed. Difficulties in being more specific about the tone and activity of some genres stem from the wide range of individuals and organisations using them with different objectives. So, for example, while an individual microblog user might use their messages to playfully communicate with friends, a business might use theirs to advertise to potential customers, and a club might use theirs to inform members about upcoming events.

Medium factors

Tables 17 and 18 offer a comparison of the primary web genres in which Irish language interaction was found, according to Herring's (2007) ten medium factors.

Genre	**Synchronicity**	**Message transmission**	**Persistence**	**Message size**	**Communication channels**
Blogs	Asynchronous	One-way	Persistent	Effectively limitless	Text, images, audio, video
Comments on news/magazine websites	Asynchronous	One-way	Varied	Varied	Text
Boards/Forums	Asynchronous	One-way	Persistent	Effectively limitless	Text, images
Listservs	Asynchronous	One-way	Persistent	Effectively limitless	Text
Microblogs	Asynchronous	One-way	Semi-persistent	140 characters	Text
Social networking sites	Asynchronous	One-way	Persistent	Effectively limitless	Text, images, audio, video
Wikis	Asynchronous	One-way	Persistent	Effectively limitless	Text

Table 17. Irish web genres according to medium factors 1-5.

Synchronicity. Computer-mediated communication (CMC) is synchronous when individuals must be online and logged on simultaneously in order to interact (Herring, 2007). This describes the type of communication that takes place through chat rooms, instant messaging services and voice-over-internet services like Skype. All of the web genres found attracting significant numbers of Irish language users were asynchronous; that is, individuals could leave messages on blogs, discussion forums, social networking sites, etc., which could be responded to by other individuals hours or days afterwards. There were no chat rooms found using the Irish language. In the past the magazine website Beo! provided an Irish language chat room, but this feature has since been removed from the website, presumably because of lack of activity. Now the only interaction that takes place on that site is through readers' comments on articles posted there. Use of Irish via instant messaging or voice-over-internet services like Skype was not found in the study. These web services are most commonly used in private networks of individuals and, therefore, their conversations would not commonly be publicly accessible.

Message transmission. All of the web genres where Irish language interaction was found involved one-way message transmission. Individuals could not receive feedback on their messages until they had finished writing them and had posted them online. This is to be expected of asynchronous CMC. Two-way transmission describes the process of simultaneous feedback which is available on some forms of CMC, through voice-over-internet or through some split-screen instant messaging services. However, no such examples were found among Irish language users.

Persistence. Persistence refers to how long messages remain online after they have been posted. Some genres, for example blogs, save all posted messages and comments allowing users to search back through an archive to the very first post. Messages left on other genres, such as chat rooms or instant messaging services, disappear after a certain time, after a specific period of activity or once the user has logged off. All of the genres popular for Irish language interaction were persistent or semi-persistent, and individuals could read back through and comment on old content. The microblogging service Twitter is semi-persistent: users can scroll back through the

3,200 most recent tweets on each account although this can be time-consuming.[78] The persistence of comments left on news/magazine websites depends on how long articles are kept online. Beo!, Nós and Nuacht24 archived all of their articles and comments back to their first editions.

Message size. Most of the web genres popular with Irish language users do not impose restrictive message buffers. The word limits on messages written in comment forms and social networking sites, for example, may vary from site to site, but in general message buffers were not seen to overly impinge on individuals fully expressing themselves. The exception was in the microblog Twitter, where participants are restricted to 140 characters per message. This study will later show how this restriction affects Irish language discourse on the site.

Communication channels. This refers to the various types of media that users can share across the different genres. Three of the genres – comments, listservs and microblogs – allowed text only, although within this text users often included hypertext links to images, audio or video elsewhere on the web. Commenters and microblog users also often included an avatar with their messages, in the form of a thumbnail image. Two genres – discussion forums and wikis – primarily used text and images. Images, here, refers to both photographs and graphics. Again, hypertext links to audio and videos on external websites could be included. In the case of Wikipedia, individuals could post images and audio relevant to the topic to which they were contributing. However, interaction on the 'talk' pages was through text only. Finally, blogs and social networking sites included text, images, audio and video in their content. It was common for bloggers to share photographs from their localities in their posts. Irish language audio and video was rare, though bloggers sometimes embedded YouTube videos of Irish language singers or television programmes in their posts. One blogger posted regular videos of his own progress learning Irish.

[78] For a more detailed description of Twitter's persistence, see http://www.ehow.com/info_12024211_long-tweets-last.html.

Genre	Anonymity	Private messaging	Filtering	Quoting	Message format
Blogs	Potentially anonymous	Not possible	Possible	Manual	Reverse chronological posts, chronological comments
Comments on news/magazine websites	Potentially anonymous	Not possible	Via mediator	Varied	Chronological
Boards/Forums	Potentially anonymous	Not possible	Via mediator	Varied	Reverse chronological threads, chronological comments
Listservs	Usually non-anonymous	Possible via individual email	Possible	Automatically quotes previous emails	Reverse chronological
Microblogs	Potentially anonymous	Possible via Direct Messages on Twitter	Possible	Through retweeting on Twitter	Reverse chronological
Social networking sites	Usually non-anonymous	Possible	Possible	Manually	Reverse chronological updates, chronological comments
Wikis	Potentially anonymous	Not possible	Via mediators	Manually	Chronological

Table 18. Irish web genres according to medium factors 6-10.

Anonymity. All web genres were found to have the potential for anonymous participation. This was more likely to occur in some genres than others. So, while listserv messages tended to include a sign-off using the participant's real name; messages on discussion forums, wikis and comment forms tended to use a pseudonym, albeit frequently with personal information about the individual in their profiles. Blogs and microblogs were split between those using their real names and identities, and those published through a pseudonym. Again, some personal information was generally provided in the users' profiles. Social networking sites like Facebook, in contrast, tended to use real names only. This practice makes it easier for users to find their friends and family online. Because access is usually restricted to real-life acquaintances on social networking sites, there is less risk of users' personal information being abused.

Private messaging. A specific function for private messaging was only available on microblogging and social networking sites. This provided a means for friends and

followers on Facebook and Twitter, for example, to contact each other privately away from the public messages posted to their walls or public feeds. There was no specific function for listserv users to contact each other individually. However, the personal email addresses included in their messages could be used to initiate more private correspondence. In other web genres, in particular blogs, users often included a personal email address in their profiles, and could be contacted privately via email. This was not compulsory, however. Private messages remained beyond the reach of this study.

Filtering. Many of the web genres attracting Irish language interaction allow users to filter out unwanted messages and to block unwelcome users. This is most obvious on social networking sites where users need to be confirmed as an individual's 'friend' before they can post messages on their homepage. Microbloggers can block other users from directly messaging them; bloggers can choose to filter comments to their posts and to not approve unwanted messages; listserv participants can block messages from individual email addresses. Individual members of discussion forums, commenters on news/magazine websites, and contributors to wikis cannot directly filter or block specific users. However, administrators of these sites can decide to filter messages and to block users if their behaviour is deemed inappropriate.

Quoting. Some websites offer a function to directly quote another individual's content so that responses can be more focused. This feature was seen on some discussion forums and comment forms, for example. In emails from listserv members, responses automatically included the previous emails to which they were responding, unless the user decided to delete them. In the case of blogs and wikis, users could choose to respond to a post in general, or to reply specifically to a previous comment. This would not quote the previous comment, but would structure the thread in such a way as to make clear to whom the comments were addressed. In most cases, however, if blog users wanted to specify the comment or the part of a comment to which they were responding, they addressed the individual in the opening of their comment, or copied and pasted a segment of the comment into their response. Quotes in this context were often preceded by the @ sign and followed by the individual's response. Facebook users could choose to 'like' individual comments, but could only reply directly to the original post in a thread. Quoting in this context was not the norm.

Twitter users have a unique way of quoting previous messages by 're-tweeting' them. This effectively attributes the message to its original author. If Twitter users want to add their own comment to the original message, there is a practice of writing their (short) responses at the start of the new message before the letters RT.

Message formatting. This refers to the order in which threads of messages appear on screen. Content in most genres was published reverse chronologically. The most recent microblog messages always appear on top of the screen. Likewise, in a thread of listserv emails, the most recent email appears on top of the correspondence. Other genres where the most recent message was prioritised include blog posts, social networking site updates, and discussion board and wiki threads. Subsequent comments on this content, however, appeared chronologically in a thread of conversation beginning with the earliest comment. Threads on some blogs, discussion boards and wikis could be interrupted if respondents chose to reply to a particular comment from earlier in the thread. Comments made to news/magazine websites also appeared in chronological order.

Situation factors

Tables 19 and 20 offer a comparison of web genres according to Herring's (2007) eight situation factors.

Participation structure. The majority of Irish language interaction was found on public websites where individuals published their material to a wide audience in a one-to-many structure. Exceptions to this structure were found among microblog and social networking site users. Although the vast majority of Irish language Twitter accounts analysed were public, there were some instances of users who kept their content private; hidden from those users who were not accepted as 'followers' of that account. Most individual Facebook accounts were kept private, with only a small amount of information about the user visible to non-friends. This was not the case with Facebook pages owned by organisations or companies, which were generally open to the public. Open Facebook groups (like the *Gaeilge Amháin* group analysed later in the study) were also publicly viewable.

Genre	Participation structure	Participant characteristics	Purpose	Topic or theme
Blogs	Public, one-to-many	Varied	Varied, information-sharing, social	Varied, personal
Comments on news/magazine websites	Public, one-to-many	Varied	Varied, expressing opinion	Varied
Boards/Forums	Public, one-to-many	Varied	Varied, expressing opinion, knowledge transfer	Varied
Listservs	Usually subscribers only, one-to-many	Varied	Varied, information-sharing, knowledge transfer	Varied
Microblogs	Usually public, one-to-many	Varied	Varied, social, information-sharing	Varied, personal
Social networking sites	Usually friends only, one-to-many	Varied	Varied, social	Varied, personal
Wikis	Public, one-to-many	Experts on topics	Information-sharing, knowledge transfer	Varied

Table 19. Irish web genres according to situation factors 1-4.

Participant characteristics. Participants varied across the genres in age, gender and location. Some genres attracted a younger audience than others, with Twitter in particular being popular among younger users. It was noted, however, that young children of primary or early secondary school age were not identified interacting in Irish on any of these genres. This would appear to chime with the findings from Fleming and Debski's (2007) study of Irish schoolchildren, where the vast majority of children in English medium, Irish medium and Gaeltacht schools claimed never to log onto Irish language websites, and rarely sent emails or text messages in the language.

Certain blogs posted content specific to a particular place or region, e.g. the Múscraí Gaeltacht in Co. Cork, and it can be assumed that they were more attractive to contributors from those regions. In general, however, there was a mix of participants from across the Gaeltacht, Ireland and the globe. Some websites attracted certain types of participants based on the topics being discussed. For example, the Daltaí discussion board often discussed Irish language grammar and vocabulary and, therefore, attracted people involved in teaching and translating the language. It can be

assumed that contributors to wikis were experts, or at least knowledgeable, in their topics of contribution.

Purpose. Purpose refers to both the group purpose, or *raison d'être*, of each website, as well as the "goals of interaction" held by their members (Herring, 2007). The reasons for participating in each genre were many and varied, and members' motivations for participating were as individual as the members themselves. Some websites had clearly stated goals, e.g. political discussion on the forum Politics.ie,[79] but others were less prescribed in the type of interaction expected from their members. Despite the varied reasons for participating across individual websites, some trends across the genres can be noted. Microblogs and social networking sites were particularly sociable, with members commonly sharing day-to-day thoughts and musings. This was also a common feature of journal-type blogs, with individuals describing and responding to stories of everyday life. Other blogs, such as those based on themes of politics, literature, or blogs maintained by community groups, were more focused on information sharing between members and the more formal expression of opinions on specific topics.

In comparison, participation in discussion forums and listservs was commonly motivated by fact finding: posing a question, problem or position to other participants and encouraging them to share their knowledge or opinions on the topic. This was particularly the case with groups discussing the Irish language itself. For example, on the Daltaí forum Irish language users of all levels discussed grammar and vocabulary, often in response to posts asking for help with translation. Interaction in wikis was primarily based around problem solving: discussing the merits of some contributions, correcting mistakes, encouraging other members to make contributions. Comment forms on news/magazine websites enabled readers to express their opinions on individual articles and to directly address the authors and other readers on a given topic.

Topic or theme. Topics varied widely across genres and websites. These were explicit in some cases, e.g. the Irish Gaelic Translator 'translation' forum,[80] or assumed from the website within which the interaction was hosted, e.g. discussions about education

[79] http://www.politics.ie
[80] http://www.irishgaelictranslator.com

on the Irish language discussion forum hosted by the Second Level Support Service for teachers. In other cases, interaction varied widely according to the particular topics that inspired individuals to compose blog posts, wiki entries, discussion board threads and listserv emails. The Irish language itself was a frequent topic of conversation. Microblogs and social networking sites also exhibited a wide range of topics. However, these genres were most popularly used to discuss the day-to-day activities and musings of their users.

Genre	Tone	Activity	Norms	Code
Blogs	Varied according to topic	Varied, information exchange, phatic exchange	Links with other blogs through blogroll	Fully formed sentences
Comments on news/magazine websites	Varied according to topic	Varied, expressing opinion	Often provocative, especially when anonymous	Fully formed sentences
Boards/Forums	Varied according to topic	Varied, expressing opinion, information exchange	Often provocative depending on topic, flaming	Fully formed sentences, with frequent emoticons
Listservs	Varied according to topic, usually formal	Varied, information exchange, expressing opinion	Often written in letter structure, with an opening greeting and end sign-off	Fully formed sentences
Microblogs	Usually informal, playful	Varied, phatic exchange, information exchange	Twitter: Use of # to group topics; RT to denote retweets; DM to send direct messages	Abridged sentences, innovative ways of maintaining message buffer
Social networking sites	Usually informal, playful	Varied, phatic exchange	Access restricted to 'friends', 'like' button	Abridged sentences, creative use of language, emoticons
Wikis	Varied according to topic	Information exchange	Joint enterprise in developing knowledge	Fully formed sentences

Table 20. Irish web genres according to situation factors 5-8

Tone. This describes the manner or spirit in which communication is made (Herring, 2007). Like topic, this varied across genre. However, there was a general distinction between the casual sociability of microblogs and social networking sites, compared to the more formal information sharing and opinion expression on wikis, listservs, news websites and some discussion forums. That said, some Twitter feeds and social networking groups were set up by groups and organisations purely to inform readers,

of upcoming events for example; and much of the interaction that took place on the other genres, seeking help with translation for example, used informal conversational features associated with friendship and familiarity. Blogs could be found across the spectrum, with some focusing on more formal discussions of news and current affairs; while others were used as a form of online journal, with all the casual language and informality one might associate with writing a diary, albeit one that could be read by a wide audience. Moreover, some blogs mixed between formal and informal conversations as the author switched between topics with each new post.

Activity. This refers to the types of exchanges taking place between individuals. It can be loosely divided between those interactions involving phatic exchange, information exchange, and/or opinion expression. Individuals might easily switch between these three functions in one conversation, let alone one genre. A spectrum of activity across the sample – from information to phatic exchange – sees wikis positioned at one end, with microblogs and social networking sites at the other. The interaction taking place on the 'talk' pages of Wikipedia was generally task-orientated with individuals sharing information with other contributors about changes or corrections to the site. The listservs in this study, as well as the language and translation-themed discussion forums, also focused on information exchange between members, providing a place where they could seek responses to problems, and to promote events that might interest other users. Individuals often used news/magazine websites and blogs to express their opinions on a given topic, in formal or informal ways depending on the topic and tone. Finally, phatic exchanges involving routine contact about everyday life was prevalent in microblogs and social networking sites.

Norms. Norms of behaviour varied across websites and genres. As individuals get used to using specific websites and through regular contact become familiar with other users, different norms of behaviour are developed. Each web genre deserves its own ethnographic analysis to better understand the culture and routines that exist among their regular users. That task falls outside the scope of this study.

Code. Irish language interaction is the focus of this study. The websites included in the study used modern Irish in text-based CMC, with all sources written in modern

Latin script.[81] As a general rule, the acute accent (or *síneadh fada* in Irish) common in Irish orthography was included. Herring (2001: 617) describes how web users innovate their writing styles online to deliberately economise on writing effort. This is one of the reasons, for example, that abbreviations are common in CMC. In this environment, it might be expected that diacritics would frequently be excluded, particularly in those genres characterised by more spontaneous responses, e.g. microblogs and social networking sites. However, even in these genres the use of the acute accent was widely adhered to. This accent indicates the length of vowel sounds, and its exclusion would risk confusing similarly spelt words with widely different meanings. For example, the words "sean", "séan" and "Seán" are all spelt with the same letters, but the different positions (or non-use) of the accent differentiates between the adjective "old", the verb "deny", and the boy's name "Seán" respectively. Kevin Scannell describes an alternative way of writing the acute accent that was used in Irish language listserv groups before the advent of 8-bit email in the 1990s (Scannell, 2007: 4). Users would insert a forward slash after the accented vowel, so that the word "béal" would appear as "be/al", for example. However, this practice was not observed in this study of contemporary Irish language CMC.

Code-switching frequently took place, in particular between Irish and English. This occurred in a number of different ways from direct insertion of non-Irish words, to bilingual translations, to the use of hypertext links from difficult words to their definitions.

Common features of Irish language online discourse

The genre analysis identified some features of CMC that appear to be particularly common in sites of Irish language interaction online, and they deserve further discussion. Firstly, all of the web genres in which regular interaction was found were examples of *asynchronous* communication. Asynchronous CMC may have particular benefits for minority languages such as Irish. Joinson (2003) explores some of the psychological conditions of internet behaviour, arguing that asynchronous text-based communication frees participants from the constraints of face-to-face interaction. Because asynchronous CMC removes the pressure to reply instantly, users can devote cognitive resources to composing their messages online away from distractions, and

[81] There was one interesting example of a blog, created by the author, that allows the reader to switch between modern Latin and traditional Gaelic type. See http://www.cainteoir.com.

spend time editing their messages. In the context of Irish, this may liberate learners or non-fluent speakers of the language to engage in interaction online without the pressure of spontaneous response. Messages may be deciphered and responded to at the readers' leisure, allowing them to consult dictionaries or grammar guides if necessary. Irish language blog authors in particular often aided readers in this process by creating hypertext links in their posts from what they perceived to be difficult words or specialist terms to their online dictionary definitions. Secondly, because asynchronous CMC does not require participants to be online simultaneously, it enables speakers located in different time zones to interact. For a language like Irish, asynchronous forms of CMC empower users around the globe to plug into a minority language community without being located in the language's geographical heartland. Many blog, Twitter and Facebook users were scattered around the globe, with a particularly large contribution from the United States. These networked individuals are afforded the opportunity to participate regularly in a network of Irish language users that might otherwise be difficult to initiate and impossible to maintain.

Irish language CMC is also largely *persistent*, with content posted to a wide range of genres remaining online indefinitely, often in the form of searchable archives. The persistence of these messages creates a source of Irish language CMC that can be analysed to see how the language is used in a contemporary computer-mediated context. Kevin Scannell of Saint Louis University, Missouri is involved in a number of projects that analyse CMC in a range of minority languages. His *Crúbadán* project, for example, uses a web crawler to recognise and collate a growing corpus of online content in Irish and other languages (Scannell, 2007). Such corpora have been used to develop online tools, to help web users with spelling and grammar, for example, or to aggregate user content from a shared language.

All of the genres included in this study could be used *anonymously*, with none requiring users to identify their real names. In spite of this, it was common for users across these genres to use their personal names (or versions of them) as usernames, in signing off messages and comments, or in their user profiles. Bloggers, Facebook users and microbloggers frequently addressed each other directly using their personal names. Some users included photographs of themselves as avatars, especially on social networking sites, and less so on blogs and microblogs. Others included background information about their locations, jobs and families in their user profiles. It has been found that anonymity on the web may lead to an increase in anti-social

behaviour (Donath, 1999). Anti-social or aggressive behaviour was not prevalent in Irish language interaction, however. Among those participants that remained anonymous there were some examples of aggressive comments in blogs and news/magazine websites, particularly associated with political debates, but these were very much in the minority. This might be a result of the relatively small number of websites with Irish language activity. With such a small potential audience, abusing other members could risk isolation. Instead the language appeared to form a bond between users, who might potentially have little contact in the language other than that shared through the web.

Designated and undesignated spaces

Sites of Irish language interaction online can be classified into two main categories spanning the various genres. Loosely speaking, these two categories represent the official face of Irish language online services in contrast with the more spontaneous private networks of Irish speakers online. Both categories differ in their purpose for participation.

Designated spaces. This refers to those websites, or parts of websites, that are specifically set up for an Irish language audience. These websites are most often provided by Irish language media companies, educational groups or third sector organisations, and are commonly supported by government funding. They include the Irish language news/magazine websites that encourage readers to comment on their articles, as well as the discussion forums set up by language groups, educational groups and others, to encourage Irish speakers to engage with each other on particular topics. They often have a clearly stated group purpose, frequently making their aims clear to users before they enter. For example, one listserv in the study (Gaeilge-A at http://listserv.heanet.ie) stated in Irish that it was for "conversation in Irish Gaelic". Similarly, a designated Irish language forum on Politics.ie is described in Irish as being for "political debates, conversation about anything in Irish". The websites are designed and administered centrally, and it is the webmasters' decision what material appears and what activity is filtered or which users are blocked.

This structure is particularly well suited to the genres of discussion forums and news/magazine websites, where from the user's point of view, it is clear that they are entering an Irish language enclave of the web. Once they have navigated to one of

these sites, they know that Irish is accepted as the main language of interaction and communicate accordingly. They can expect that other Irish users with similar interests will also be participating.

Undesignated spaces. This describes the popular social media services within which Irish language users have set up their own accounts. Here, it is the users who decide who to interact with and what language to use. This is the case with blogs, Twitter and personal Facebook accounts, where Irish language users might occupy just a tiny fraction of the overall activity taking place. Unlike the discussion forums, listservs and news/magazine websites, it is difficult to see how much Irish language activity is taking place across the blogosphere and on international sites like Twitter or Facebook. On these services each individual user builds up their own networks of friends, which are often hidden from public view. Users are free to upload whatever content they wish, in whatever language they wish. They may communicate entirely in Irish, or they may do so with only a small number of their online acquaintances. Among these users there is no group purpose, rather a motivation to casually interact with other Irish language users as they desire.

Blogs, microblogs and social networking sites are not just for personal use. This study found many examples of blogs, Twitter and Facebook accounts run by third sector organisations; traditional Irish language media companies; Irish language community groups; schools and university departments; and Irish language companies, such as publishers. Unlike standalone news/magazine websites or discussion boards, these are embedded into wider networks of users. These sites play an important role in linking individual users who might otherwise be unaware of each other's presence.

Within the large networks of blog, Twitter and Facebook accounts there are examples of users coming together to encourage Irish language use across these genres. Together their individual motivations for using these sites have been temporarily supplanted by a group purpose to network web users or to engage non-users in the language. This has been done by the users themselves, in some cases with help from third sector organisations. We have already discussed the forming of the *Gaeilge Amháin* Facebook group and the Indigenous Tweets website, both of which bring together Irish users in different ways. Other examples, both online and offline, include:

'Irish Day' on Twitter.[82] Beginning in 2010, a special celebration of Irish has been marked on Twitter as *Lá na Gaeilge* (Irish day) on St. Patrick's day (March 17th) each year. This was begun as a project between friends online. Irish language Twitter users are invited to upload tweets in Irish on that day including the hashtag #lnag (short for Lá na Gaeilge) in their messages. The #lnag tag identifies who's using Irish that day and creates a single thread of all the Irish language messages with that tag as they are uploaded.

The Irish blog symposium.[83] In 2011, as part of an Irish language literature festival, a symposium of Irish language bloggers was organised in Dublin by a small number of Irish language social media enthusiasts. Three bloggers and one Facebook group administrator spoke at the event – one via Skype, and other Irish language bloggers attended the event and blogged about it afterwards. The event, which was the first of its kind, discussed such issues as the blogosphere as a source of literature, the future of the blogosphere, particularly in light of the huge rise in popularity of Twitter, and the place of minority languages on the web.

The #Gtuít tweetup.[84] A 'tweetup' is a social event that brings Twitter users together face-to-face in the real world. The word is a portmanteau of the words 'Twitter' and 'meet up'. The #Gtuít tweetup was organised by the Irish language organisation Gaelport to encourage Irish language users of Twitter and other social media to come together in Dublin in September 2012. Although primarily a social event, the tweetup invited attendees to discuss a range of topics relating to the language online, including recent developments in technology and the success of online Irish language campaigns.

Having identified and analysed the most popular web genres for Irish language interaction, the rest of this research project focussed on three distinct groups from three web genres: blogs, the microblogging service Twitter and the social networking site Facebook. The study now turns to these groups, beginning with an analysis of

[82] For more information, see: http://lnag.org.
[83] For more information (in Irish), see: http://www.cityarts.ie/events/2011/10/14.
[84] For more information, see: http://www.gaelport.com/Gtuit-tweetup

each group's demographics, describing the numbers and types of people using these social media and their levels of activity.

VI THE IRISH LANGUAGE IN SOCIAL MEDIA

This study chose as its focus the three most significant online Irish language groups in terms of public activity – the Irish language blogosphere, Twittersphere, and the *Gaeilge Amháin* Facebook group. These groups were chosen because of the relatively large numbers of participants found interacting through Irish on a regular basis. Moreover, they represent three distinct genres popular in the Irish language and three of the most popular social media worldwide. Each group would be explored from the point of view of the users, and the ways in which they interacted through Irish. Participants would be defined and described via a range of research tools using both social network analysis and computer-mediated discourse analysis techniques. Furthermore, the discourse produced on-screen by each group's core members would be analysed and described to shed light on the distinct features of Irish language CMC.

A. GROUP MEMBERSHIP & ACTIVITY

Each group was analysed demographically to measure the size of the active membership, to describe the types of users participating, and to quantify participants' levels of activity. Although the three groups in this study were chosen for their relatively large numbers of participants in an Irish language context, the population of each group was small. In summary, the three groups can be described as follows:

- Each group comprised between 150-300 participants.
- Participants were more likely to be male, with at least twice as many male members than female in each group, when gender could be identified.
- Each group was found to have participants located around Ireland and the world.
- However, most participants in each group were located in Ireland, outside the Gaeltacht.
- The majority of members of each group had joined in the previous three years (the *GA* Facebook group was created in 2011).
- The majority of members in each group participated infrequently, with most messages and comments coming from a small number of prolific users.

- Topics varied greatly between users.
- However, everyday activities and musings from users' personal lives tended to be the most popular topic of conversation in each group.

Each group will now be described in greater individual detail.

1. THE IRISH LANGUAGE BLOGOSPHERE

Arguably the most accessible network of Irish language users online is that of the Irish language blogosphere (ILB). The ILB, as it is imagined in this study, is comprised of 73 blogs that had published Irish language posts between January and March 2011; the 68 individual bloggers who wrote posts on these blogs; and the 83 commenters. All of these blogs could be viewed by the wider public without the need for registering or logging in. These figures include the researcher and his blog. Each blog in this study is unique, focussing on its own topics of interest and published in the bloggers' individual writing styles. Yet, by interacting with other bloggers and commenters through writing comments, posting hypertext links, linking through blogrolls, etc., each standalone blog represents a node in a wider network, many of which are tied together in meaningful ways through regular interaction.

There may have been other active Irish language blogs during this period that remained hidden from the researcher – most likely isolated blogs with low levels of activity. However, six months after the end of the data capture period the Indigenous Blogs website began compiling data on blogs written in Irish, and only one additional blog was found to be active during the data capture period. This had attracted no comments during that period, nor did it link to any other blogs in its blogroll. It was added retrospectively to the ILB. Thus, although this study cannot assert to be a complete list of Irish language blogs, it can confidently claim to cover the majority of active Irish language blogs over a three month period in early 2011.

Describing the ILB

Gender. Tables 21 and 22 show the breakdown of blog, blogger and commenter gender. The difference in the numbers of blogs and bloggers is due to some individuals authoring more than one blog, and some blogs being authored by more

than one individual. It would appear that blogging in the Irish language is more popular among men than women. Where individuals' gender could be identified, nearly three quarters of both bloggers and commenters were male. This is consistent with Cunliffe & Honeycutt's study of Welsh language blogs (although commenters were not included in their analysis).

Gender	Male	Female	Mixed	Unknown
Blogs	41	12	4	16

Table 21. Blog gender.

Gender	Male	Female	Unknown
Bloggers	37	13	18
Commenters	42	15	26
Total	79	28	44

Table 22. Blogger/commenter gender.

Age. Age data was not collected for any of the groups in this study, as it was unavailable for most members. This contrasts with Cunliffe & Honeycutt's study of Welsh language blogs where age data was collected, albeit for less than a third of subjects.

Location	**Gaeltacht**	**Ireland, possibly Gaeltacht**	**Ireland, Outside Gaeltacht**	**Rest of World**	**Unknown**
Blogs	6	5	33	15	14

Table 23. Blog location.

Location	**Gaeltacht**	**Ireland, possibly Gaeltacht**	**Ireland, Outside Gaeltacht**	**Rest of World**	**Unknown**
Bloggers	3	5	29	14	17
Commenters	1	7	5	15	55
Total	4	12	34	29	72

Table 24. Blogger/commenter location.

Location. Tables 23 and 24 show the geographical location of each blog, blogger and commenter. There is evidence that members of the ILB are scattered across Ireland and the globe. Where specific locations could be identified, bloggers and commenters in Ireland were most likely to be located in counties Dublin (17), Cork (8), Galway (6) and Donegal (4). These correspond to the two counties with the highest

populations and largest Gaeltacht areas respectively. They are also the four counties with the highest numbers of daily Irish speakers inside and outside the education system according to the 2011 census. The most common locations of overseas bloggers and commenters were the USA (15) and Canada (4). This would reflect traditionally popular destinations for Irish emigrants. It should be noted, however, that other destinations with significant numbers of Irish immigrants, such as Great Britain and Australia, did not show up in any large numbers. Individuals were identified in other countries across Europe – in Germany, France, Italy, Estonia, Kosovo – and the world – Philippines, Brazil – possibly reflecting the destinations of more recent emigrants.

Blog characteristics

Some general characteristics of users were recorded in line with Cunliffe & Honeycutt's study.

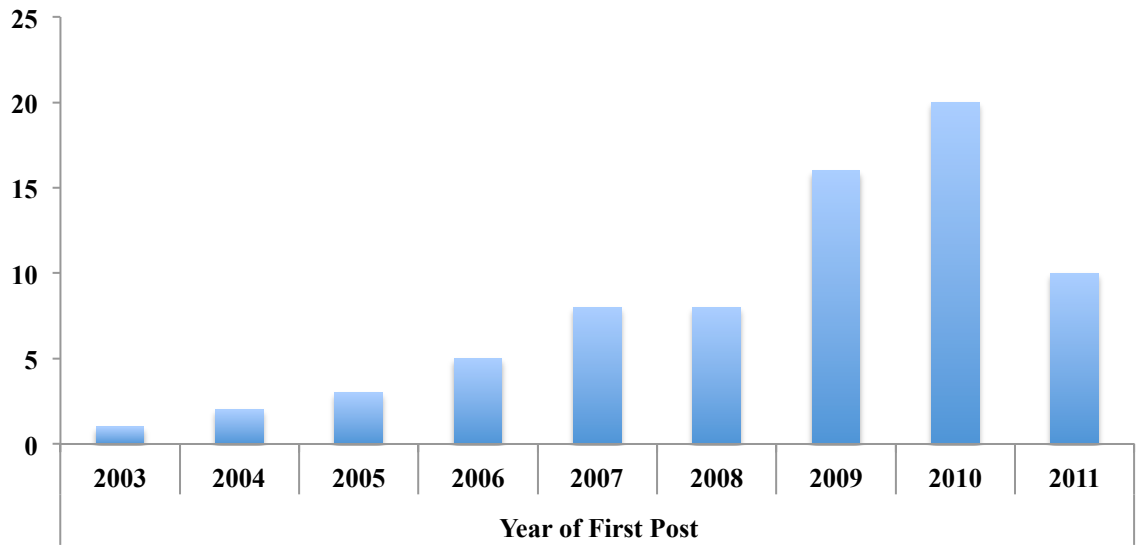


Fig. 12. Blog start dates.

Figures for 2011 pertain to Jan-Mar only

Longevity. Figure 12 shows the start dates of the 73 blogs in the ILB, as deduced from the dates of their first posts. The figure for 2011 is lower than expected as the data capture period ended in March of that year. The data shows a rise in the number of new blogs in the ILB over the past eight years. However, this does not necessarily indicate an increase in the popularity of Irish blogging since 2003; but might instead

reflect blogging fatigue and a low level of persistence, with older blogs becoming dormant and eventually dying. This low level of persistence appears to be borne out in the posting levels of older blogs. Of the six blogs that began in 2003-2005, only one posted more than five times in the Irish language during the data capture period in 2011. This contrasts with the same three month period in 2009 where three of the six blogs posted more than five times in Irish and earlier still in 2007 where four of the six blogs did so. This rather crude measure would suggest a certain amount of blogger fatigue. Despite the growth in popularity of other online genres, in particular the spike in interest in Facebook and Twitter in recent years, blogging is clearly still a popular form of online communication, with ten new blogs commencing during the three month data capture period alone.

Main topic	**Secondary topic**	**Combined topics**
Personal blogs (13)	Current Events & News (6)	Personal blogs (17)
Community blogs (12)	Literature & Poetry (5)	Community blogs (12)
Irish language (as topic) (7)	Personal Blogs (4)	Literature & Poetry (12)
Literature & Poetry (7)	Politics & Government (4)	Politics & Government (11)
Politics & Government (7)	Arts & Entertainment (2)	Current Events & News (9)
Arts & Entertainment (5)	Eclectic (2)	Irish language (as topic) (9)
Eclectic (5)	Education (2)	Arts & Entertainment (7)
Current Events & News (3)	Irish language (as topic) (2)	Eclectic (7)
Education (3)	Lifestyle (2)	Education (5)
Internet & Software (3)	Sports (2)	Sports (4)
Sports (2)	Design & Photography (1)	Internet & Software (3)
Design & Photography (1)	Food (1)	Design & Photography (2)
Food (1)	Humour (1)	Food (2)
Travel (1)	Religion & Faith (1)	Lifestyle (2)
		Travel (1)
		Humour (1)
		Religion & Faith (1)

Table 25. Blog topic.

Topic. Assigning a topic to a blog is problematic, and relies on the researcher to interpret and categorise the subject matter of each post. This study used Cunliffe & Honeycutt's informal approach to assigning topics to blogs, with a small number of adjustments. A new category "Literature & Poetry" was used to distinguish the relatively large interest around this topic in the ILB. Presumably, this would previously have fallen under Cunliffe & Honeycutt's "Arts & Entertainment" and "Fiction" categories. Finally, a new category of "Food" was created. A number of Cunliffe & Honeycutt's topics were not popular in the ILB, including "Blog Resources", "Business & Professional", "Counter Culture", "Genealogy" and "Science & Technology".

As in the Welsh language blogosphere, the most frequent topic in the ILB was that of personal blogs, with day-to-day events and musings published in a journal style making up a large part of the output of 17 blogs. There were 12 community blogs found publishing information for people interested in various groups, from poetry enthusiasts to teachers and parents of children in Irish language nurseries. Other blogs were published on a wide range of topics, with a particularly keen interest in current affairs. The strong showing of politics in particular is likely to have been helped by the general election which took place in the Republic of Ireland during the data capture period. The arts in general inspired a lot of content, with topics related to literature, poetry, music and media popular on many blogs.

Connections across the ILB. The study identified and traced 390 blogroll entries, 59 permalinks, over 2,000 comments, 115 follows, three notes, and five likes linking bloggers and commenters in the ILB. No valid trackbacks were found. Comments were by far the most frequent type of interaction across the ILB, and would later form the focus of discourse analysis.

Blog activity

This study also recorded levels of user activity.

Posts. Figure 13 shows the number of posts made by blogs in the ILB during the data capture period. As ten new blogs were commenced during this period, the figures probably slightly under-represent levels of posting activity. Across the ILB, on average 13.3 new posts were made on each blog over the three months. All but one

blog made fewer than one post per day; 48 blogs (65.8%) made fewer than one post per week; and 35 blogs (47.9%) made fewer than one post per fortnight. The two most frequent posters – GFG[85] and NI – posted 98 and 88 times respectively.

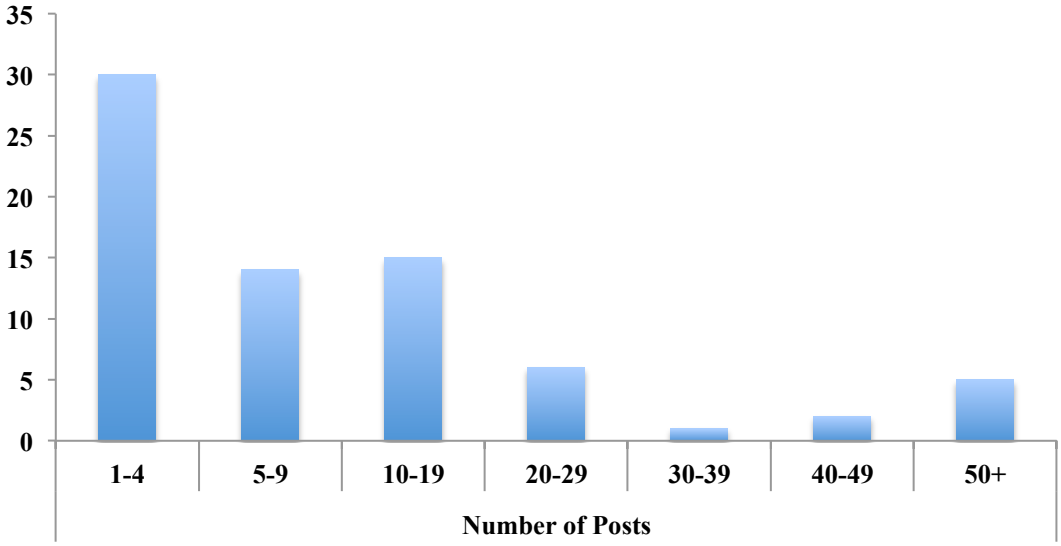


Fig. 13. Blog posts.

Comments. Figure 14 shows the number of comments made by bloggers and commenters across the ILB during the data capture period. These measures differ from Cunliffe & Honeycutt's methodology in two key ways. Firstly, data is collected for both bloggers and commenters. Rather than measuring the number of comments left on each blog (i.e. the destination of the comments), this data shows the number of comments produced by each blogger/commenter (i.e. the source of the comments) regardless of where they were written. Secondly, this study uses a different method for counting comments. Cunliffe & Honeycutt counted each comment as one unit, regardless of how many people were addressed in the message. As the focus of this study is to understand the ILB as an interactive social network, a more flexible approach was used to counting comments in an effort to shed more light on who was interacting with whom within the ILB. In many cases, within one comment, a series of communications were made responding to multiple other comments in the thread. Thus, each comment was read through and counted according to whom the blogger/commenter was addressing or responding. This was easier to identify in some instances, according to the practices by which different individuals were addressed.

[85] All blog, Twitter and Facebook users were given acronyms to protect their anonymity.

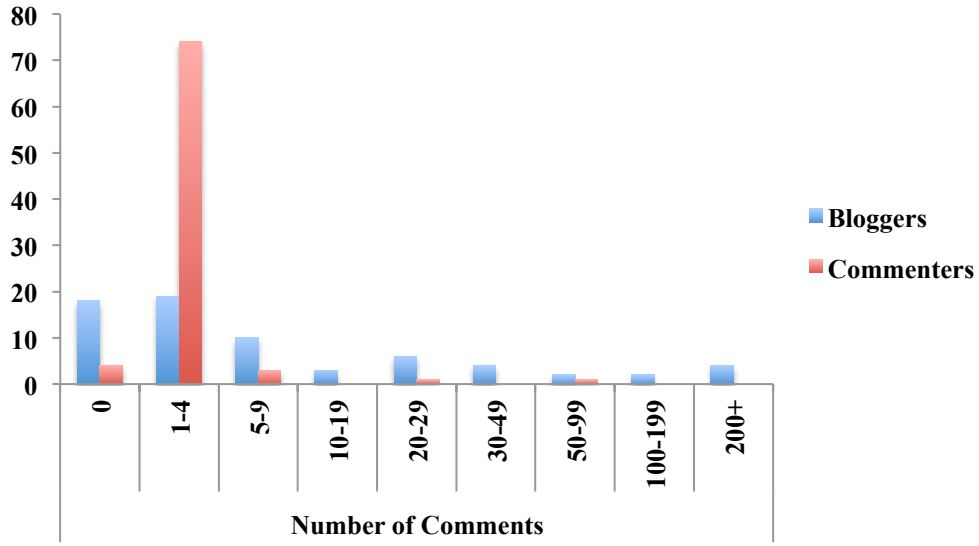


Fig. 14. Blog comments.

Over the three months each individual made an average of 18.7 comments. This does not mean they commented on blogs on over 18 different occasions, but rather that they used comments to address other individuals in the ILB on average 18.7 times. Most comments were short. 22 members of the ILB did not make any comments during the study period, but rather blogged in isolation or used other methods to interact. Many commenters only made a small number of comments during the three-month period, with ten commenters (14.7%) making only one comment and five (7.6%) commenting only twice. These represent very light users who, rather than browsing the blogosphere regularly, may have been attracted by one particular post, perhaps navigating to the post through hypertext links elsewhere on the web. On the other end of the spectrum, eight bloggers and one commenter made over 50 comments each, three of whom commented 400 times or more.

Thread length. Due to the different methodology in counting comments, analysis of thread lengths in the three groups in this study was not carried out.

2. THE IRISH LANGUAGE TWITTERSPHERE

The ILT is the most active public network of Irish language users online, in terms of messages sent. In this study it is comprised of the 274 Irish language accounts that

were active in the month of February 2012. The names of these accounts were sourced from the Indigenous Tweets website, and to be deemed active they must have posted at least one Irish language tweet during the data capture period. They must also have written at least 2.5% of all of their tweets in Irish, according to the language data on Indigenous Tweets.

Describing the ILT

Gender	Male	Female	Unknown
Users	119	63	92
Percentage	43%	23%	34%
Percentage identified	65%	35%	

Table 26. Twitter user gender.

Gender. Table 26 shows the breakdown of gender in the ILT. It would appear that tweeting in the Irish language is more popular among men than women by a ratio of 2:1, albeit with a large number of accounts uncategorised. The gender of users posting to 92 accounts could not be determined. This was because users did not share their gender in their profiles, or because people posting to the accounts of clubs or organisations could not be identified and categorised.

Location	Gaeltacht	Ireland, possibly Gaeltacht	Ireland, Outside Gaeltacht	Rest of World	Unknown
Users	36	58	138	26	16
Percentage	13%	21%	50.5%	9.5%	6%
Percentage identified	14%	22.5%	53.5%	10%	

Table 27. Twitter user location.

Location. There is evidence that members of the ILT are scattered across Ireland and the globe (see Table 27). The majority of users identified their location in their profiles, albeit to varying degrees of specificity. Where multiple locations were mentioned in users' profiles their tweets were examined to identify where they were located during the data capture period. By far the most common locations were in Ireland. The results show a majority of users – 53.5% of those identified – living in Ireland outside the Gaeltacht. A further 22.5% of identified users mentioned Ireland or a Gaeltacht county as their location without specifying whether they lived inside or

outside the Gaeltacht. Many of these are likely to live in towns and cities outside the Gaeltacht boundaries. When specific locations are counted including those profiles where more than one location is mentioned, Dublin city and county was the single most common location (71), followed by Galway city and county (40; 27 of whom specify Connemara) and Belfast (24). Most Irish language Twitter users live in urban locations, with 37% of those whose locations were identified having ties to just two cities – Dublin and Belfast. The most common overseas locations mentioned in profiles were Great Britain (11) and the USA (11); both being traditionally popular destinations for Irish emigrants.

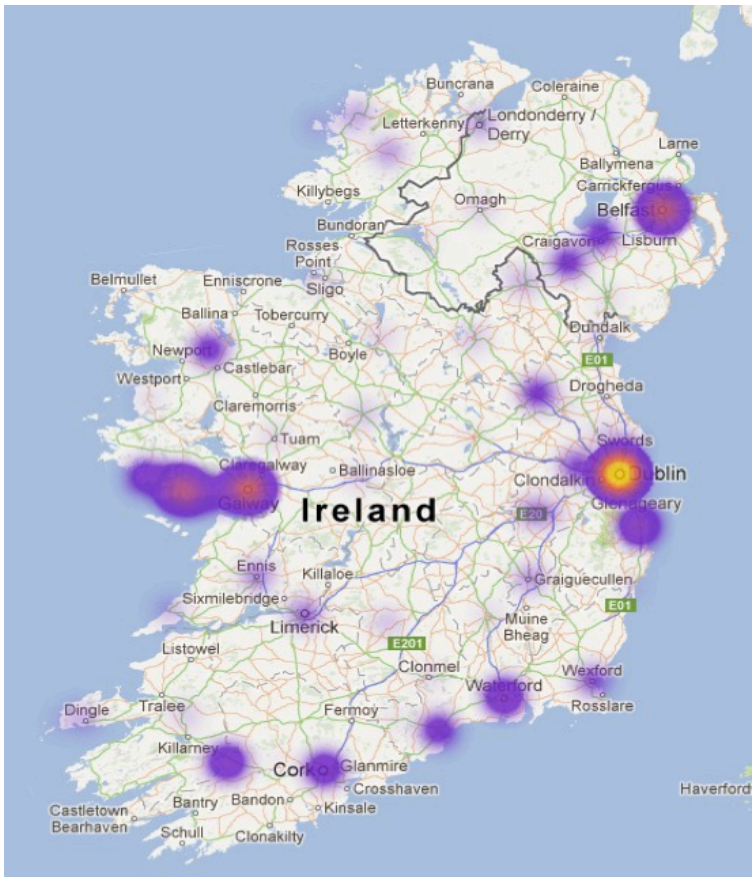
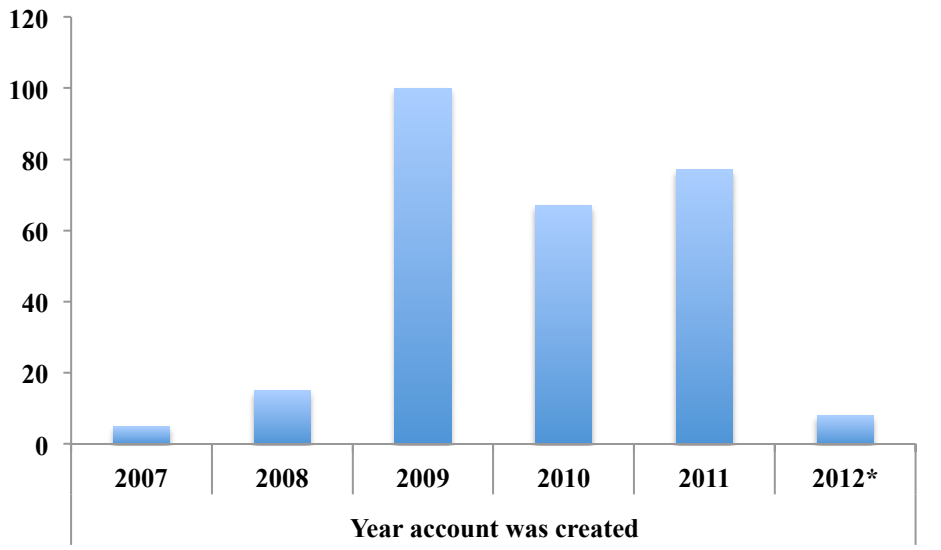


Fig. 15. Primary locations of Irish language Twitter users (Scannell, 2013).

The findings of this study matched those of Scannell (2013), where he plotted the geographic locations of all of the Irish language Twitter accounts listed on his website Indigenous Tweets. Scannell's map of the data (see figure 15) shows concentrations of users in Ireland, in Dublin, Galway/Connemara and Belfast corresponding to the most popular locations of members of the ILT in this study.

Twitter account characteristics

Longevity. Figure 16 shows the start dates of 272 accounts in the ILT.[86] The figure for 2012 is lower than expected as the data accounts for the first two months of that year only. The data shows massive growth in the number of new accounts in 2009. Although Twitter was launched in 2006 it took some time for it to grow in popularity outside the USA. Just 20 accounts active in the Irish language during the data study period were commenced pre-2009. This data shows that 120 accounts in the ILT (44% of the accounts whose start dates could be identified) had been tweeting for at least two years. It is unknown whether they had been tweeting persistently in Irish during this time, however. These users are likely to be familiar with the conventions of communicating in Twitter, and are most likely to have adapted their communication to the specific requirements of Twitter over time.



* Jan-Feb only

Fig. 16. Twitter account start dates.

Topic. Assigning topics to Twitter accounts was problematic, and relied on the researcher to interpret and categorise the most common subject matter in each case. In devising a list of categories, this study was guided by Cunliffe & Honeycutt's

[86] The start dates of two accounts could not be identified.

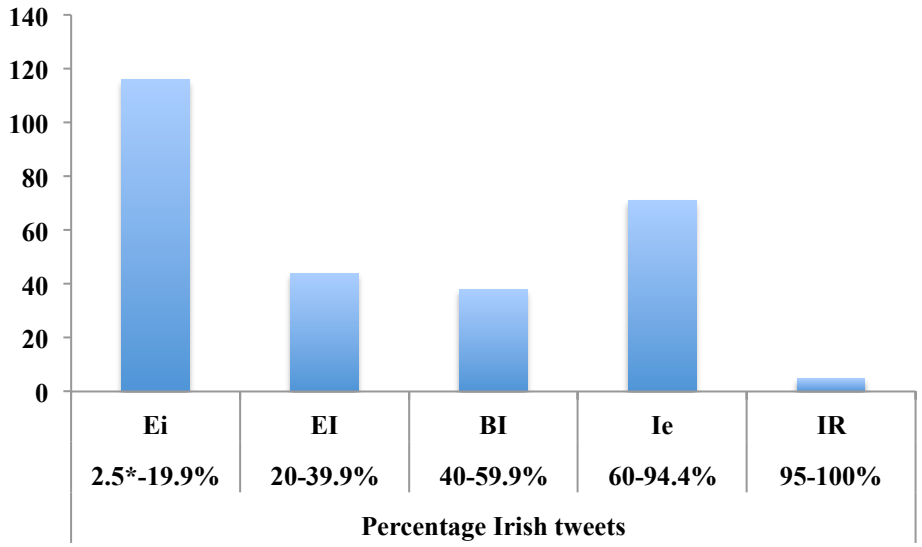
informal approach to assigning topics to blogs. Most Twitter use was more conversational in tone than in blogs, where users have more space to present and develop their thoughts on their topic of choice. For example, although there were many accounts in the ILT run by politicians in the Republic of Ireland and Northern Ireland, few of these tweeted regularly about politics itself. Instead they were observed to use the medium to communicate informally with their followers and with those who had directly messaged them. Twitter is a very social medium. By far the most common category of conversation was 'personal'. This describes those users who posted regular musings and updates on their day-to-day lives and maintained contact with their friends. Other common topics included: arts and entertainment, current events and news, education, politics and sports.

Followers and following. Data on the numbers of followers attracted by each account in the ILT, as well as the number of others they were following, was obtained directly from Indigenous Tweets. The average number of followers across the ILT was 594. The average number of accounts they followed was lower at 481. It is impossible to know whether or not followers were attracted to accounts in the ILT through their Irish language content. As most accounts in the ILT tweeted in Irish less than 100% of the time, their followers may well be engaged exclusively in their English (or other) language content.

It might be expected that the longer accounts had been tweeting, the more followers they would attract. This was not necessarily the case, however. For example, eight of the top 20 accounts by number of followers had been tweeting for less than two years. There were a number of well-known names in the February sample, including an international sports personality (boxer Bernard Dunne), several TV and radio presenters (Daithí Ó Sé, Evanna Ní Chuilinn) and a well-known journalist (Harry McGee). These tended to attract a much higher number of followers than they followed themselves, e.g. Bernard Dunne was following almost 70 times fewer accounts than were following him. When the top five celebrity accounts are removed from the sample, the ratio of average number of followers to followees is reduced to 482:480. This would seem to endorse the general link symmetry that defines online social networks in the work of Mislove *et al* (2007).

Language. Figure 17 shows a breakdown of the percentage of Irish language tweets

across members of the ILT, as listed on Indigenous Tweets. According to the data, 42% of accounts use Irish in less than one in five of their tweets. This represents a large section of the ILT who are only occasionally engaged in communicating through Irish. Around 13.9% of members tweeted bilingually (that is between 40-60% of the time in Irish). On the other end of the spectrum, more than a quarter of accounts tweeted primarily in Irish (that is, over 60% of the time), with five members (less than 2% of the ILT) composing over 95% of their tweets in Irish. This latter group represents a small number of what might be termed monolingual Irish language Twitter users.



* accounts with less than 2.5% Irish tweets had already been removed from the sample

Fig. 17. Percentage Irish language tweets.

Twitter activity

Tweets per month. Indigenous Tweets lists the total number of tweets published by Irish language accounts, as well as the number of tweets they wrote in Irish. These figures relate to all of the tweets posted since Indigenous Tweets identified each account as using Irish. This study used the data gathered on each account's start date to quantify the average number of tweets posted each month in all languages and in the Irish language. The results show that on average members of the ILT posted approximately 165 tweets per month, with approximately 29 of these written in the Irish language. The average number of Irish language tweets per month ranged from 801 to fewer than 3. The top thirty most prolific ILT members posted more than twice

per day in the Irish language. On the other end of the spectrum, the 25 least active accounts tweeted less than once a week in the Irish language.

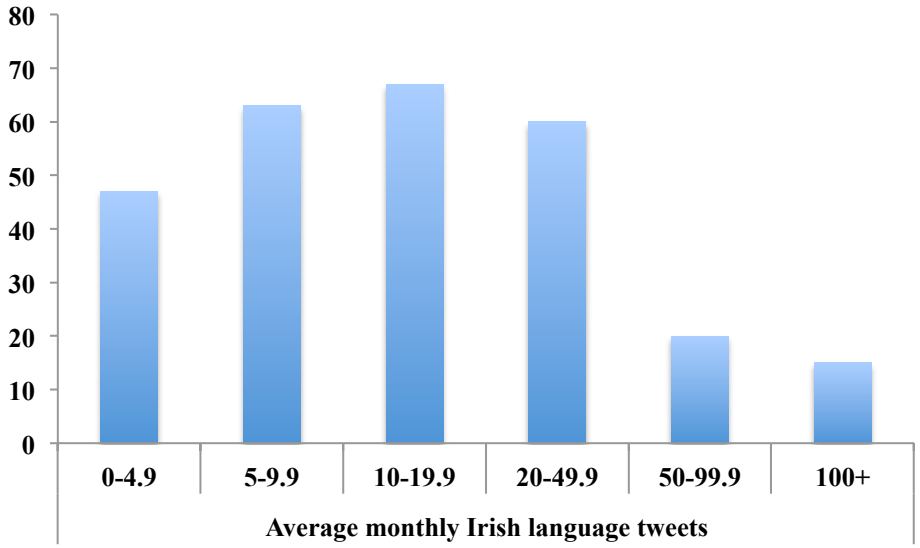


Fig. 18. Irish language tweets per month.

3. THE *GAEILGE AMHÁIN* FACEBOOK GROUP

The Irish language Facebook Group *Gaeilge Amháin* differs from the ILB and the ILT in that all interactions take place on one group page (or wall), where all members can read the most recent posts and comments. The group was created in 2011 with a mission to encourage Facebook users to interact through Irish only. Since then it has attracted over 1,600 members with various levels of activity. The *GA* network as it is imagined here consists of the 265 individuals who were found actively participating in the group in August 2012.

Describing the GA membership

Gender. Table 28 shows the breakdown of *GA* members' gender. Activity in the group is more popular among men than women: 65% of members who could be categorised.

Gender	Male	Female	Mixed	Unknown
GA members	158	85	1	21

Table 28. Gender in *GA*.

Location. Tables 29 & 30 show the geographic location of *GA* members, indicating that membership is scattered across Ireland and the globe. Unlike the ILT and ILB, however, membership appears to be concentrated in one particular region. Of the 174 members identified as living in Ireland both inside and outside the Gaeltacht, 79 of them lived in the province of Ulster in the north of the island. The group founder – Ailéin Ó Clumháin – lives in Northern Ireland, and many of the most prolific contributors are located in counties Donegal, Derry and Antrim. There is a particularly strong link to the Gaeltacht areas of Co. Donegal, and many members appear to have attended Irish language courses there in the past. Moreover, many of the photographs of group members posted to the site come from social outings in Co. Donegal and neighbouring counties. In late 2012 the group's cover photo displayed at the top of the homepage was changed to a photograph of Mount Errigal, an iconic mountain in Co. Donegal, thereby further cementing the group's strong identity with the region.

Location	Gaeltacht	Ireland, possibly Gaeltacht	Ireland, Outside Gaeltacht	Rest of World	Unknown
GA members	18	12	94	31	110

Table 29. *GA* member location.

Location	Gaeltacht	Possibly Gaeltacht	Outside Gaeltacht
Ulster	9	3	67
Non-Ulster	9	9	27

Table 30. Locations of *GA* members in Ireland.

Member characteristics

Longevity. Facebook publicly reveals some limited information about when members joined each open group. Figure 19 shows the start dates of 259 of the *GA* members still active in August 2012.[87] The majority of active *GA* members had joined before October 2011 and were still contributing posts and comments to the group wall. Figure 19 indicates that the group consistently attracts small numbers of users each month who remain active over time. Twenty new members had joined during the data capture period, the highest rate of active users from the previous eleven months.

[87] The start dates of six members could not be identified.

However, judging by the trend over time, it is expected that after an initial flurry of activity some of the new August members will cease participating in the group.

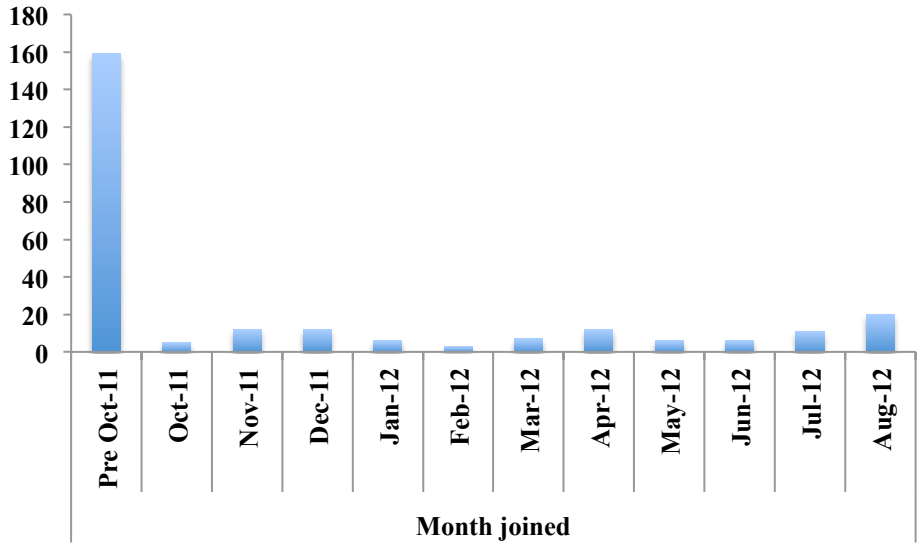


Fig. 19. Member joining dates.

Topic. Topics were not assigned to individual members of *GA*. Unlike the blogosphere and Twittersphere where individuals set up distinct accounts through which they focus on their own lives, hobbies and interests, *GA* is a shared site where members participate in group discussions on a wide range of themes and topics. Although some members preferred making certain types of posts – e.g. humourous posts, proverbs, group photos – the majority of members could not be assigned specific topics of interest.

GA activity

Posts. Figure 20 shows the number of posts made by *GA* members to the group wall in August 2012. The majority of members did not contribute posts, preferring instead to comment on other members' posts. Five members posted over 25 times, posting new messages or photos almost every day. These highly active members are particularly important, as their contributions provide new material around which other group members can interact. The five most prolific posters were all well established in the group and had been members for over eleven months by the data capture period.

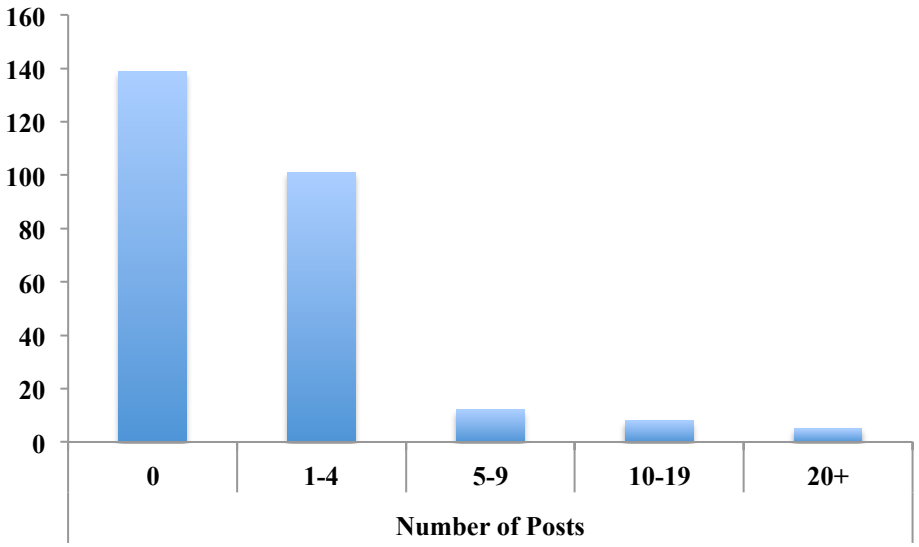


Fig. 20. *GA* posts.

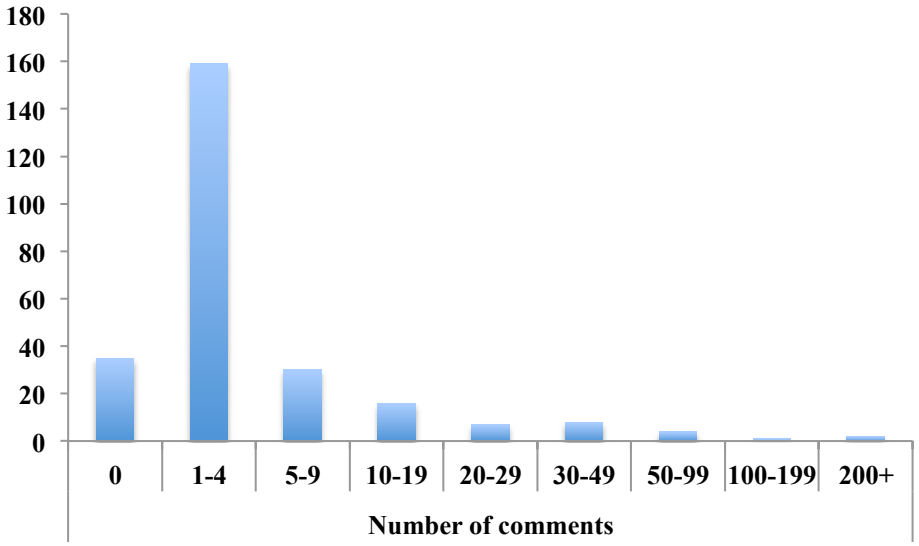


Fig. 21. *GA* comments.

Comments. Figure 21 shows the number of comments made by *GA* members in August 2012. The numbers refer to how many times each member addressed someone else in the group. As one comment might address or respond to more than one individual in the body of its message, some messages were counted as more than one comment. This method for counting comments was used to focus on *GA* as an interactive social network. 35 *GA* members did not respond to any content posted by other members, and their participation was restricted to simply publishing original

posts on the group wall. The majority of group members who did comment on other people's posts did so infrequently, with 60% of members only responding to others between one and four times over the month. However, there was a small group of ten users who commented over 50 times during the data capture period. These heavy users played a key role in engaging others to interact.

B. VISUALISING EACH NETWORK

Using relational data from each group, a social network based on users' interactions was created for each group. Each group was shown to share some basic network characteristics:

- Each group comprised one main network component.
- Each group had a number of unconnected isolates representing those participants who were not in direct contact with other members.
- Each group had a small but densely connected network core representing the most prolific users.
- Less active members in each group were loosely connected to the network and located on the periphery.
- Newcomers tended to be well integrated into each network, with some newcomers located in the networks' cores.
- Language use varied across the networks.

The network structure of each group will now be discussed separately.

1. THE IRISH LANGUAGE BLOGOSPHERE

Figure 22 shows a social network visualisation of the ILB, with individual members appearing as nodes, and with edges linking nodes that shared at least one connection across the blogosphere during the data capture period. The researcher's own blog is included in the network, although data on ties to other blogs through the researcher's blogroll was not. Because the researcher added any newly found Irish language blogs to his blogroll during the period of participant-observation, its inclusion might have overly impacted the structure of the network.

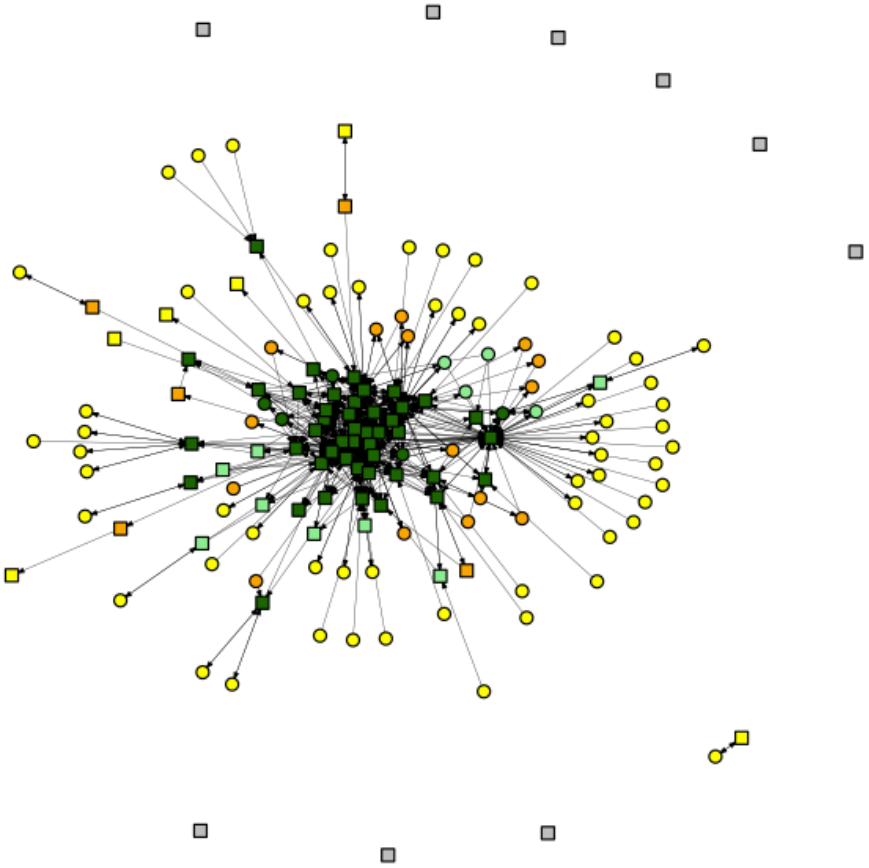


Fig. 22. Social network visualisation of the ILB.
Nodes are colour-coded according to the number of ties they share.
Dark green > 3 ties, light green = 3 ties, orange = 2 ties,
yellow = 1 tie, grey = isolate.

In this visualisation, blogs appear as squares and commenters appear as circles. It was decided to show blogs rather than bloggers in the visualisations to help identify the sites of most frequent interaction. Only four blogs are authored by multiple bloggers, so the majority of square nodes in this visualisation represent just one blogger. In six cases, where bloggers had authored more than one blog, individuals are represented by more than one node. In figure 22 nodes have been colour-coded according to the number of other ILB members they were connected to. Edges are directed, with arrows indicating the direction of the ties. The nodes are displayed using the Fruchterman-Reingold layout algorithm in the igraph package in R (Fruchterman & Reingold, 1991; Csárdi & Nepusz, 2006). Fruchterman and Reingold developed this force-directed two-dimensional layout to display nodes in an aesthetically-pleasing way, pulling together well-connected nodes, and pushing nodes

that are less well-connected to the periphery. In this way, it acts as a clear visual comparison of network density and centrality (Herring et al, 2007).

Network structure. The most striking feature of figure 22 is how interconnected the ILB is, with just nine blogs and one blog-commenter dyad isolated from the main network component. This is potentially a result of the link-trace method of finding new blogs and commenters. It is clear that through blogrolls, permalinks, comments and other connections the majority of members of the ILB were connected to other individuals in some way during the study period. Over half share a tie with just one or two others (shown in yellow and orange). These represent the ILB's poorly connected periphery. From a structural perspective, there emerges two particularly interesting points of study. Firstly, this visualisation would suggest that there is a small but densely clustered group of blogs at the centre of the ILB (colour-coded in dark green); and, secondly, that there is a small number of blogs near the centre of the ILB that are the sole focus of relatively high numbers of commenters. These might best be understood as 'authorities' in the ILB – highly-referenced nodes in the network (Gibson et al, 1998) attracting relatively high inbound compared to outbound links (Herring et al, 2005). Both the central cluster and 'authorities' would form the focus of further discourse analysis.

Language. It is significant that the ILB has a monolingually Irish core. Figure 23 indicates that all of the bloggers and commenters that were well connected at the centre of the network used only Irish in their communications. The interactions and relationships they share are done exclusively through Irish. Embedded within a wider multilingual web, they are mutually engaged in a joint enterprise to create and preserve a monolingually Irish social network. Discourse analysis would later reveal how they have adapted their CMC to that end.

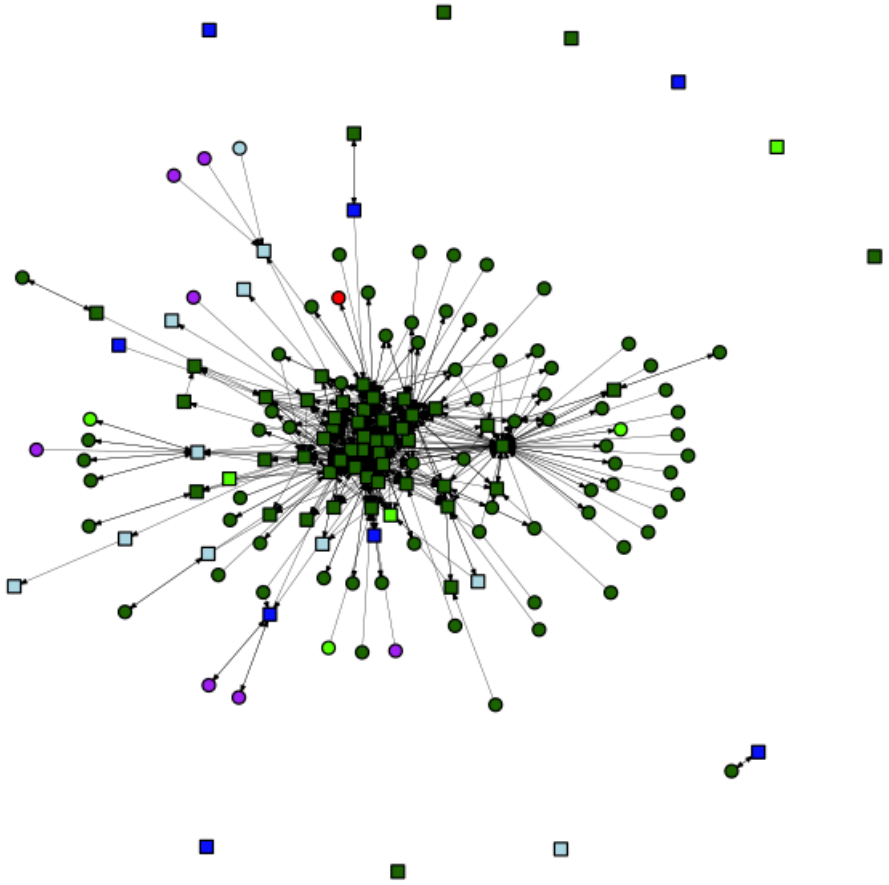


Fig. 23. Language in the ILB.
Nodes are colour-coded according to their language use.
Dark green: IR, light green: Ie, light blue: BI,
dark blue: EI, purple: EN, red: SG.

Robustness. It is clear that the ILB is a network with a relatively small population of just 151 individuals communicating across 73 blogs. This is comparable to the 75 Welsh language blogs found on one blog aggregator in Cunliffe & Honeycutt's study, but overall occupies just a tiny fraction of the millions of blogs worldwide. To put it in context, Herring *et al's* (2005) study of over 5,000 English language blogs, identified one network of approximately 130 Catholic blogs. Thus, just one small network of English language blogs linked around a shared topic outnumbers the entire Irish language blogosphere. In addition, many members of the ILB exhibit low activity levels, with half of blogs being updated less than once per fortnight. Moreover, the blog start dates show that most bloggers in the ILB are relatively recent arrivals and suggest blogger fatigue among longer established participants.

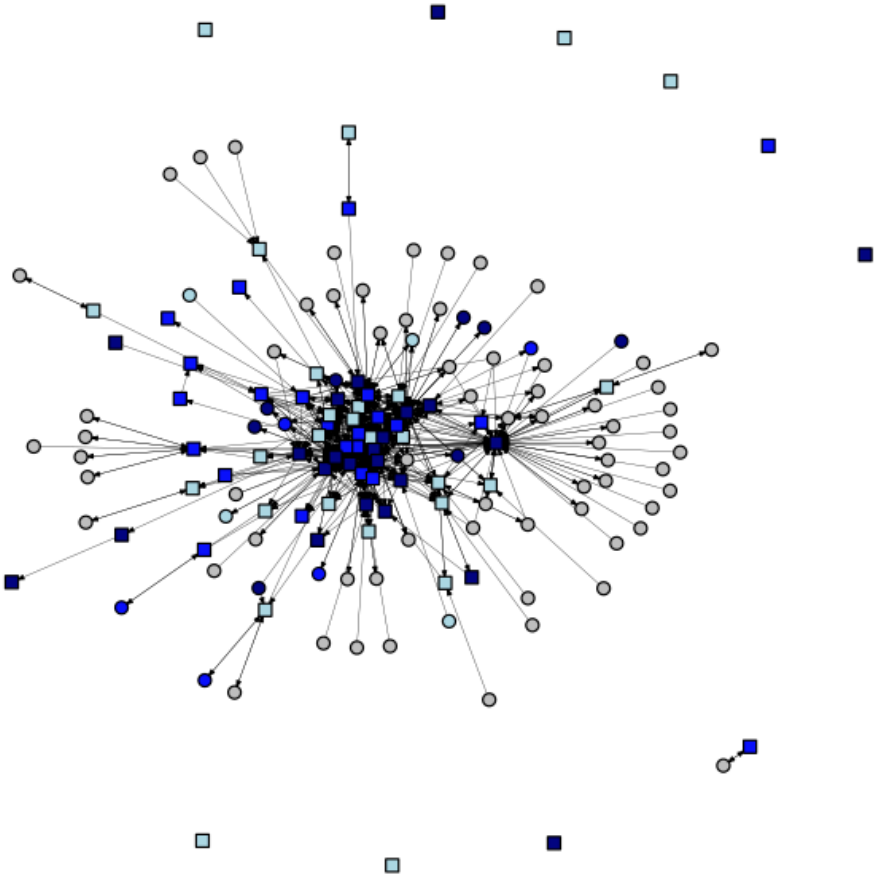


Fig. 24. Newcomers in the ILB.
Nodes are colour-coded according to when they began blogging/commenting.
Navy: 3+ years ago, blue: 1-3 years ago,
light blue: <1 year ago, grey: unknown.

The ILB does, however, exhibit some characteristics of a robust language network that augur well for its survival. Firstly, from a linguistic point of view, 123 blogs and commenters (78.8%) communicated monolingually through Irish, with a further six (3.8%) communicating mostly through Irish. Figure 23 indicates that monolingual Irish language users were concentrated in a small but dense and well-defined core, with those members of the ILB communicating bilingually or primarily through English occupying peripheral positions in the network. Secondly, it would appear from figure 24, where members are colour-coded according to their start dates, that newcomers are integrated well into the network. The central cluster is split between those who had been active in the blogosphere for over three years, and those who had joined the blogosphere in the past three years. Furthermore, there are a small number of individuals in the core who had only commenced blogging in the past year. This

suggests that if interest in blogging continues the ILB can be sustained into the future through newer members, if older members become less active in the network. With such a fluid membership, however, this process might change the character of the network over time.

In March 2013, two years after the initial study of the ILB was conducted, just 39 of the 73 Irish language blogs were still active (i.e. they had posted new content in the previous three months). There is evidence that some bloggers have migrated to Twitter, to the detriment of their blogging activity. Further analysis would be needed to confirm if enough new blogs are currently being created to replace the dead ones.

2. THE IRISH LANGUAGE TWITTERSPHERE

Figure 25 shows a social network visualisation of the ILT as it is imagined in this study. Nodes represent individual Twitter accounts from the February sample, with edges representing three or more @replies, mentions or retweets referring to other members of the network in Irish language tweets. Edges are directed, with arrows indicating the direction of messages. Since the relational data goes back to March 2011 (i.e. approximately eleven months of Twitter activity) it was decided to set a threshold of three or more @replies, mentions or retweets to indicate those nodes that had interacted directly with one another in a more than casual way during this period. Nodes without any such ties to other members of the network appear as isolates on the edge of the visualisation.

Network structure. The first striking feature of the visualisation in figure 25 is the relatively high number of isolates, with 54 nodes – almost one in five – sharing no ties with others in the main network component. These isolates may well have shared contact with others in the ILT and may be following or followed by them, but at fewer than three Irish language tweets directed to any other member of the ILT over an eleven month period this interaction was very infrequent. Many of them posted only occasionally in the Irish language and may interact more frequently with users in other languages. Moreover, this study does not include direct messages (or DMs) sent privately between Twitter users.

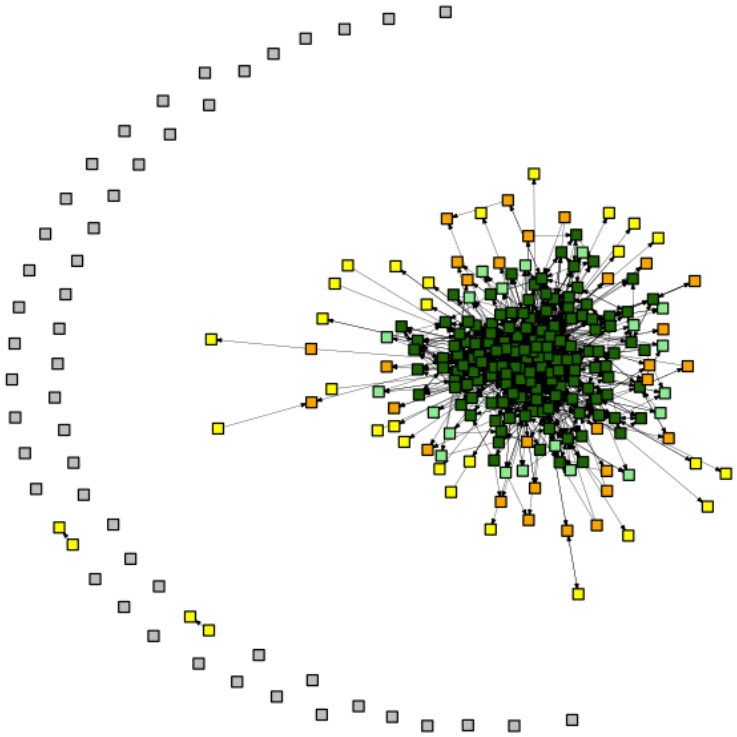


Fig. 25. Social network visualisation of the ILT.
Nodes are colour-coded according to the number of ties they share.
Dark green > 3 ties, light green = 3 ties, orange = 2 ties,
yellow = 1 tie, grey = isolate.

There do not appear to be any substantial clusters of users separated from the main network component. The lack of private clusters of users is all the more surprising given Twitter's popular use in sharing conversations and 'pointless babble' (Crystal, 2011) between acquaintances. It is possible that more tightly grouped clusters of friends may emerge from the main network component at higher thresholds of interaction.

Over a third of nodes in the main network component were only tied to one or two others. These represent a poorly connected network periphery (shown in orange and yellow). In the centre of the network there exists a dense cluster of well-connected users. These core users are in contact with multiple others through ties that are usually reciprocated.

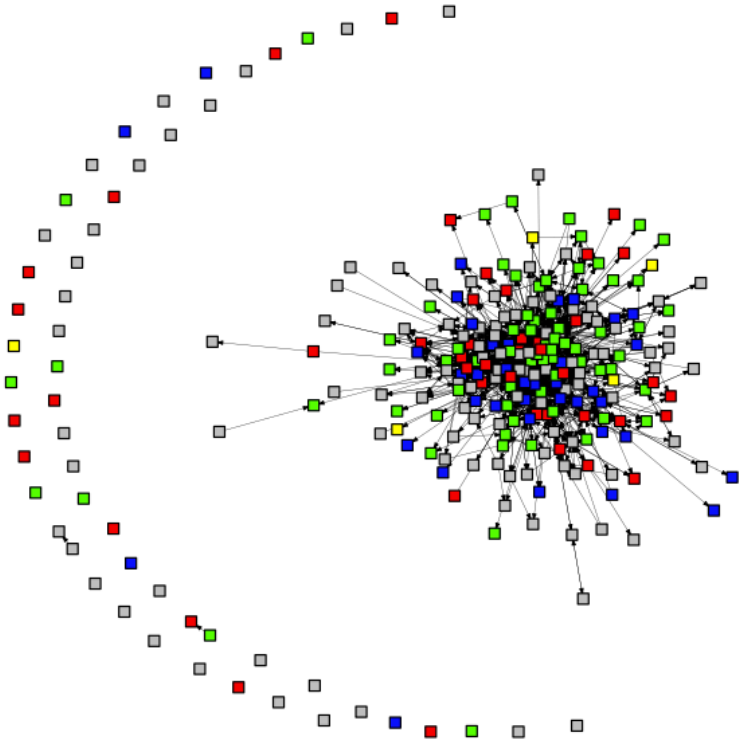


Fig. 26. Social network visualisation of the ILT.
Nodes are colour-coded according to their language use.
Yellow: IR, green: Ie, blue: BI,
red: EI, grey: Ei.

Language. Unlike the ILB, the ILT does not have a monolingual Irish core. In fact, of the five monolingual Irish nodes, all are located on the periphery. Instead, figure 26 indicates that the core consists of users with a spectrum of Irish language activity. This may have a negative impact on the development of a distinctly Irish language register in Twitter, as a large share of individuals situated at the core of the ILT experience Twitter primarily through English. For these users interacting in the Irish language occupies only a fraction of their time on Twitter.

Figure 27 displays the ILT social structure again, but this time with nodes colour-coded according to the average number of Irish language tweets they published each month. Here we can see a pattern in the structure of the ILT, whereby the most prolific Irish language participants are the best connected in the network, with the least prolific located predominantly on the network's periphery. At first glance there might seem an obvious correlation between tweeting regularly in Irish and interacting with many others in the ILT. However, referring back to figure 26 we see that the core of prolific Irish language tweeters includes many individuals who tweet (at best)

bilingually, and (at worst) in Irish in only a small fraction of their overall tweets. Twitter does not have a feature for filtering a Twitter account's output by language. Thus, if one follows an account with only a fraction of tweets in Irish, the Irish tweets risk being lost in the predominantly English (and other language) output.

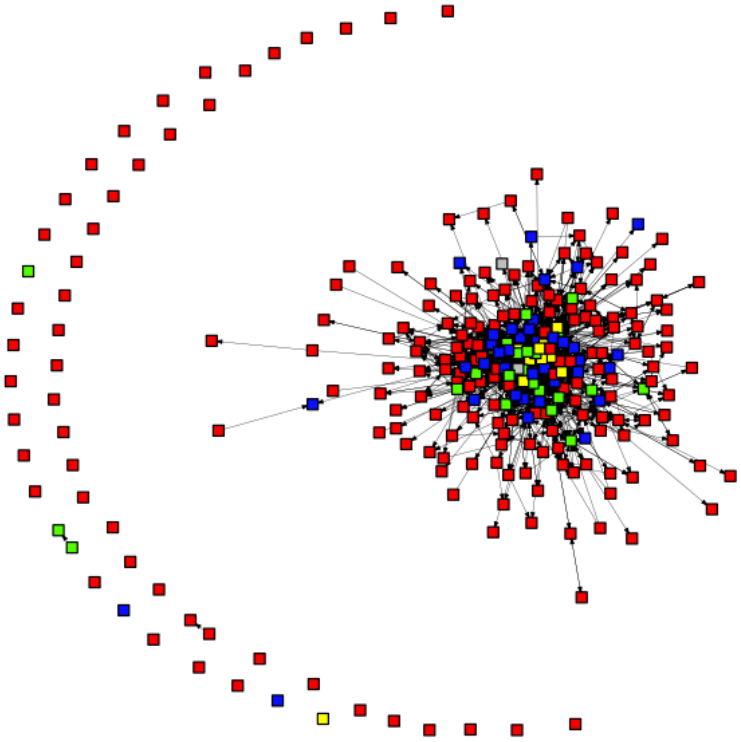


Fig. 27. Prolificity in the ILT.
Nodes are colour-coded according to their average monthly Irish tweets.
Yellow > 120 (4 per day), green > 60 (2 per day), blue > 30 (1 per day),
red < 30 (1 per day), grey = unknown.

Two features in Twitter help prevent this from happening, however. Firstly, hashtags allow the folksonomic display of tweets on various topics and themes. Searching for popular Irish language hashtags, or for common Irish words, enables users to collate content in or about the language from across the Twittersphere. This helps to promote the content of the most prolific Irish language users, while filtering out their non-Irish content. Secondly, the unique function of addressing and mentioning users directly in tweets via their @usernames enables users to focus on their individual (Irish language) conversations with others in the Twittersphere. Both @replies and mentions are compiled in separate sections of a user's Twitter account allowing them quick access to the messages that are directly related to them. This

enables users to maintain Irish language conversations with some followers while tweeting elsewhere in English and other languages. The position of prolific (but multilingual) Irish language users at the core of the network suggests that users are availing of these functions to help focus on Irish language content and conversations.

Robustness. It is clear from the demographics that the ILT, as it is imagined here, is a network with a relatively small population of just 274 individual accounts, one-fifth of whom do not interact with other Irish language accounts with any frequency. Moreover, many members exhibit low activity levels, with approximately half of accounts posting fewer than twelve Irish language tweets per month. More than half of ILT members have only been tweeting since 2010. In this context, it is difficult to predict the robustness of such a recently established network.

It would appear from figure 28 that new users are integrated well into the network's core. It indicates a mixture of Twitter experience across the core of the ILT, with many newcomers to Twitter – shown here in yellow – connected to core members of the network, and many experienced users – shown here in grey – located on the periphery. As web habits evolve and as user fatigue sets in, the group of individuals using Twitter in Irish is likely to change over time. The absence of a dense cluster of monolingual users at the core of the ILT suggests that the network might in the future be threatened by such changes. However, the successful integration of newcomers into the network augurs well for the continued survival of a network of Irish language Twitter users.

In March 2013, one year after the initial study of the ILT, 345 of the most prolific Irish language Twitter accounts listed on Indigenous Tweets were still active (i.e. they had tweeted in Irish in the previous month, and wrote 2.5% or more of their tweets in Irish). The figures indicate that tweeting in Irish is at least as popular as when the initial analysis was carried out, if not more so. This augurs well for the future of the ILT.

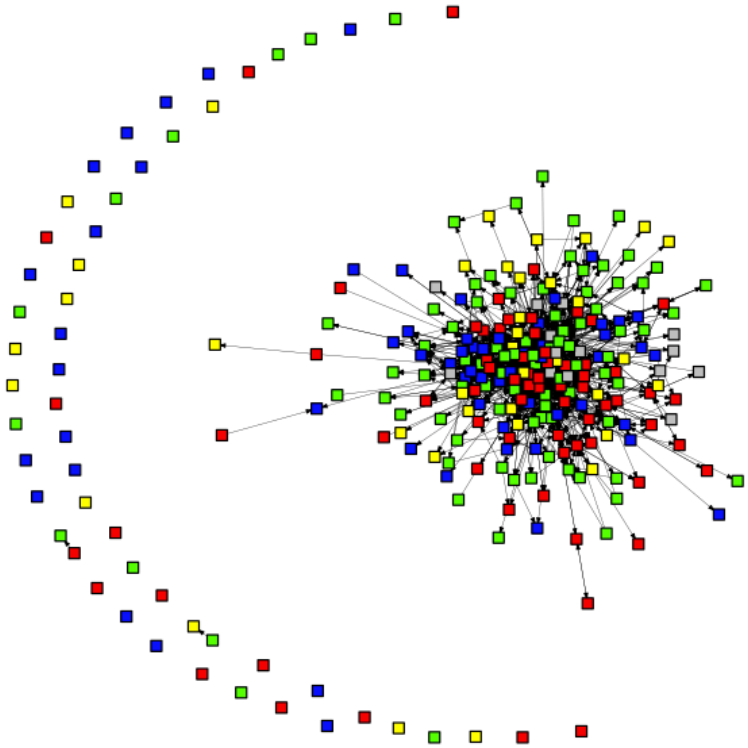


Fig. 28. Start dates in the ILT.

Nodes are colour-coded according to when their account was first registered. Yellow: July 2011-Feb 2012, green: July 2010-June 2011, blue: July 2009-June 2010, red: July 2008-June 2009, grey: pre-June 2008.

3. THE *GAEILGE AMHÁIN* FACEBOOK GROUP

Figure 29 shows a social network visualisation of the *GA* network. Each node represents one member of the group, colour-coded separately according to their number of ties. Individual nodes are connected by ties to those other members they have addressed or responded to at least once in public messages on the group wall in August 2012. This is admittedly a very low threshold to be considered actively interacting with others in the network. Ties are directed with arrows illustrating the direction of interaction.

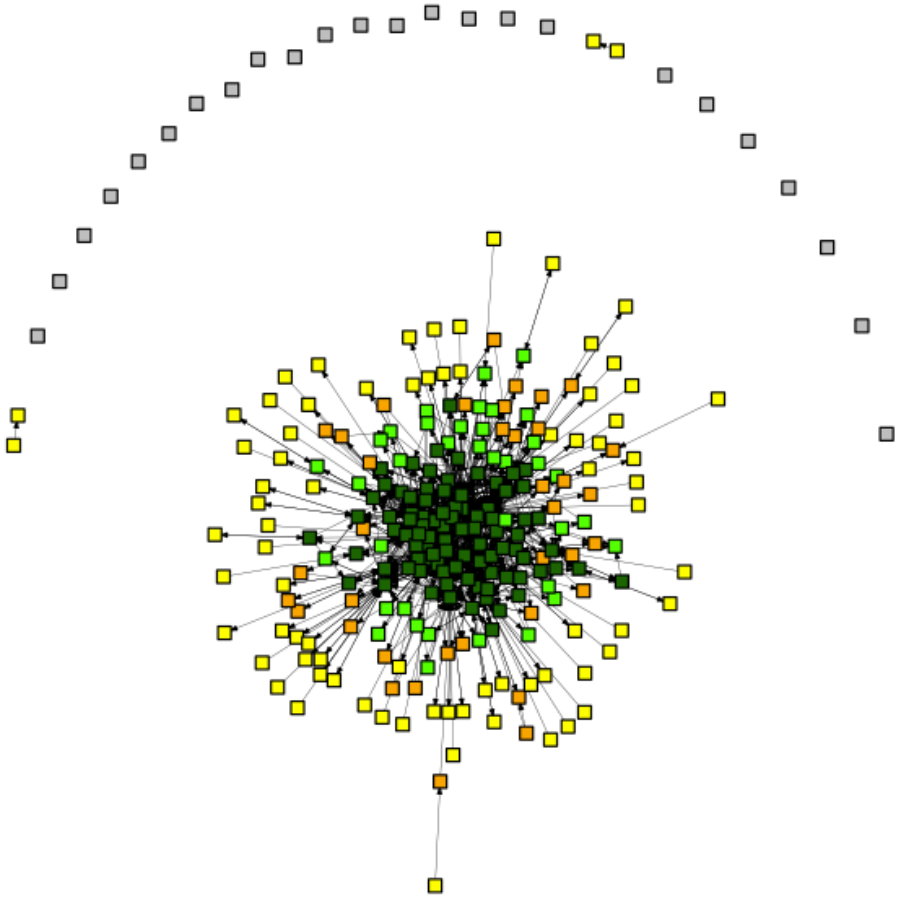


Fig. 29. Social network visualisation of *GA*.
Nodes are colour-coded according to many ties they share.
Dark green > 3 ties, light green = 3 ties, orange = 2 ties,
yellow = 1 tie, grey = isolate.

Network structure. Despite the large proportion of members connected in the main network component, the *GA* group is a relatively loosely tied network. 24 members are isolates, meaning they made at least one post during the data capture period without attracting comments or responding to other people's posts. Approximately half the nodes in the main network component are only tied to one or two others (shown in yellow and orange). At the core of the network is a small group of densely connected nodes, including the group's administrators. The content posted by these core members will be analysed later.

Language. There was no variation in language use across the network, as the group administrators operated a strict Irish only policy. All members, therefore, communicated monolingually in Irish.

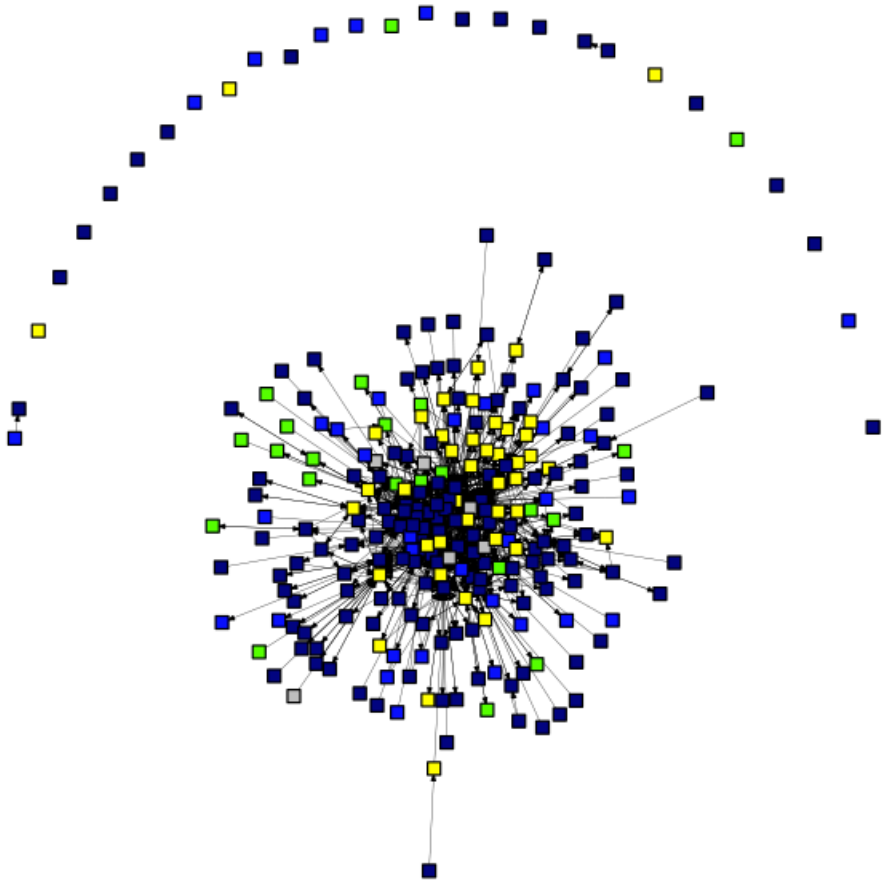


Fig. 30. Longevity in *GA*.
Nodes are colour-coded according to how long they've been members.
Navy > 11 months, blue 6-11 months, green 3-6 months, yellow < 3 months.

Robustness. It is clear from analysis that the *GA* Facebook group is a network with a relatively small and loosely tied membership. Moreover, many members exhibit low activity levels. *GA* does, however, exhibit some characteristics of a robust language network that augur well for its survival. From a linguistic point of view, all members interacted solely through Irish. This was strictly enforced by the group's administrators and established members. Regular messages reminded members of the Irish only rule, and English language content was removed from the group wall. Core members appear to be familiar with one another, addressing each other informally, and sharing photographs and stories from real life social events. This suggests that ties at the core are multiplex and shared outside of the Twittersphere. Furthermore, it would appear from figure 30, where members are colour-coded according to their start dates, that some newcomers are integrated well into the network. Although the majority of core users have been members of the group for at least eleven months,

some core members had joined in the previous three months. However, it remains to be seen how many of these newcomers persist in contributing to the group over time.

In March 2013, six months after the initial study, *GA* had over 2,100 registered members. This growth in membership augurs well for the group's continued survival. However, further analysis would be needed to determine how actively new members are using the group to interact through Irish.

The next stage of the study identified the core users in each group and conducted discourse analysis on their interactions. This would reveal how the Irish language has been adapted to each genre, and shed light on the registers peculiar to each group.

C. LANGUAGE & REGISTER

Firstly, the core users in each group were identified by measuring individuals' degree centrality in each network. Table 31 shows degree centrality measures for the best-connected blogs in the ILB in descending order. The number of comments and posts made by each node during the study period is also included.

Rank	Blog ID*	Degree Centrality	Comments	Posts
1	RAM	0.581	454	62
2	SFA	0.568	498	50
3	IG	0.497	45	41
4	GFG	0.445	400	98
5	NI	0.419	357	88
6	AOT	0.387	29	14
7	CLR	0.329	152	46
8	GA	0.323	0	23
9	ADS	0.265	12	2
10	ACA	0.252	42	2
11	RNG	0.245	137	19
12	SFI	0.219	30	26
13	CHD	0.213	68	1
14	DS	0.194	9	21
15	FC	0.187	79	23
16	ACO	0.168	5	13
17	UIE	0.148	22	11
…	…	…	…	…
154	FIS	0	0	2
155	GSB	0	0	1
156	MMM	0	0	1

*Each blog, commenter, Twitter account and *GA* member was given a unique three-letter acronym to maintain their anonymity throughout this study.

Table 31. Degree centrality in the ILB.

Rank	ID	Degree centrality	Monthly Irish tweets
1	aon	0.707	801.8
2	mur	0.414	151.3
3	iga	0.322	121.3
4	sao	0.319	220.4
5	sco	0.319	116.1
6	bre	0.282	233.7
7	mit	0.271	114.6
8	tde	0.231	46
9	the	0.190	83.9
10	dub	0.179	74.3
11	mil	0.168	98.6
12	pis	0.168	62.2
13	nei	0.161	167.6
14	gap	0.161	126.7
15	dou	0.154	116.9
16	sil	0.154	49.2
17	ohf	0.150	49.2
18	sen	0.147	39.7
19	ohe	0.143	38.8
	…	…	…
272	amh	0	3.2
273	dan	0	3
274	soi	0	3

Table 32. Degree centrality in the ILT.

Rank	ID	Degree centrality	Posts	Comments
1	MaB	0.697	39	305
2	AiC	0.636	68	257
3	ToM	0.265	25	49
4	AnK	0.261	18	99
5	BrB	0.258	28	170
6	OiM	0.22	8	88
7	AoG	0.212	19	26
8	PeM	0.208	0	86
9	SeM	0.201	15	46
10	LuM	0.186	0	55
11	FrC	0.178	29	47
12	CiD	0.174	19	37
13	BrF	0.170	2	56
14	MaG	0.159	6	50
15	EaG	0.148	13	22
16	BeB	0.144	3	43
17	PaH	0.14	8	42
18	ReL	0.125	10	42
19	SeH	0.125	7	9
	…	…		…
272	RoD	0	1	0
273	SeC	0	1	0
274	SoG	0	1	0

Table 33. Degree centrality in *GA*.

Tables 32 & 33 show equivalent data for the ILT and *GA*. It would appear from these figures that nodes with high degree centrality are also relatively prolific in each network. In interacting with multiple others, it was hoped that nodes with high degree centrality would be influenced strongly by other members of their network in how they communicated. It was assumed, therefore, that if language norms existed in each group they would be most apparent in the discourse of core members.

The top 14 best-connected blogs and the top 16 best-connected Twitter users and *GA* members were selected for linguistic analysis. Figures 31, 32 and 33 show the positions of these core nodes in each network. They confirm the strengths of the Fruchterman-Reingold layout algorithm in creating social network visualisations where well-connected nodes appear at the centre and poorly-connected nodes are pushed to the periphery.

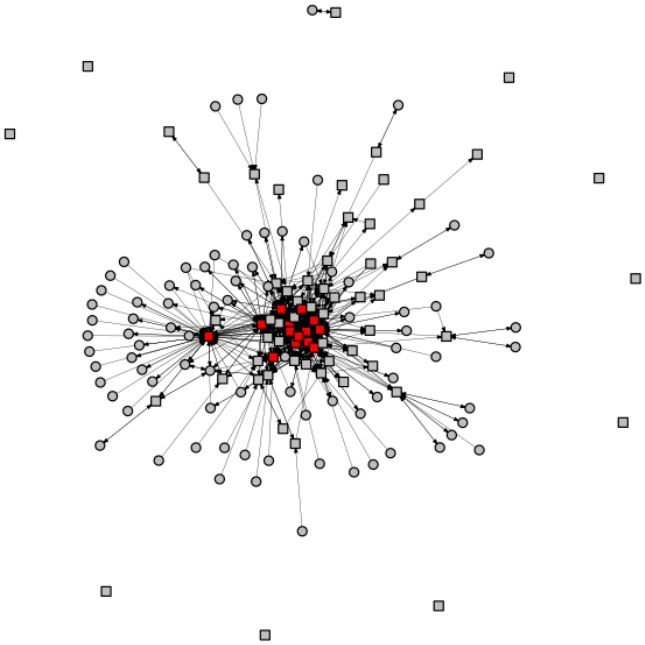


Fig. 31. Core members of the ILB.
Red nodes indicate those with the highest degree centrality.

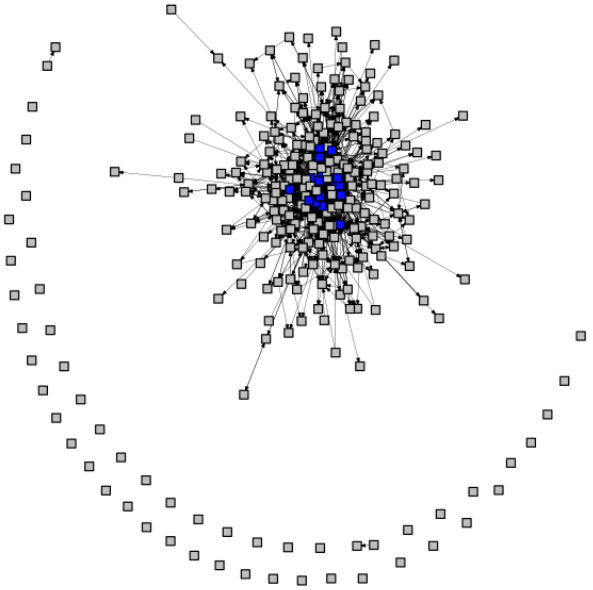


Fig. 32. Core members of the ILT.

Blue nodes indicate those with the highest degree centrality.

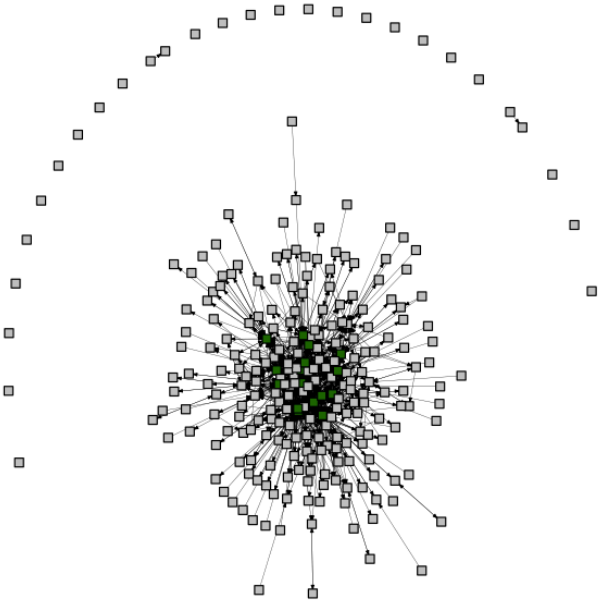


Fig. 33. Core members of *GA*.

Green nodes indicate those with the highest degree centrality.

All of the posts, comments and tweets made by these core users during the data capture periods were downloaded and compiled into corpora for linguistic analysis. This analysis specifically sought to describe the discourse in each group according to the features used by Cherny (1999) to characterise the distinct register of an MUD community (syntactic and morphological variation, abbreviations and acronyms, play

with modality, and language routines) as well as code switching practices. In general the written discourse in each group could be described as inventive, with frequent use of non-standard spellings and expressions. There was a spectrum of language use: blog posts most often comprised standardised fully-formed sentences; while tweets were most often composed of casual non-standard language and signs written in truncated sentences. The *GA* Facebook group lay somewhere in between, with frequent non-standard language use sometimes associated with low language competency among users operating in an all-Irish language environment. Each of Cherny's features, as well as code switching, will now be discussed in greater detail with examples given from each group.

1. SYNTACTIC & MORPHOLOGICAL VARIATION

There were a number of features associated with syntactic and morphological variation that were shared across the three groups. These can be summarised as follows:

- Words were commonly spelled in non-standard ways reflecting how the words sounded in the users' own speech.
- Punctuation was frequently ommitted to reduce typing space, time and effort.
- Logograms and digits were commonly used in place of written words for the same reasons.
- Occasionally in the ILT and *GA* Facebook group digits were used in words in place of syllables that would be pronounced similarly when read, in a practice common to English language text messaging and CMC.
- Despite frequent non-standard spellings, members of the ILB and *GA* Facebook group regularly corrected mistakes in their own or others' spellings and grammar.

Each feature of syntactic and morphological variation will now be described in greater detail with examples.

Syntactic and morphological variations are commonly used in online discourse, primarily to economise on typing effort, mimic spoken language features or for creative expression (Herring, 2001). They were common to all three online groups,

but especially the Twittersphere where the 140-character message buffer adds an extra incentive to users to minimise their message length. This can be compared to similar requirements in the use of SMS messages on mobile phones. In his study of a small sample of English language tweets, Crystal (2011) discovered that less than 10% exceeded the 140-character message buffer. When these were removed from his sample the average tweet length was 100 characters. Not one tweet from the corpus of 2,473 Irish language @replies, mentions and updates from ILT core users was found to exceed Twitter's message buffer. The average character count across the corpus was much lower, at 87.8 characters per tweet. Indeed, there were many examples in the corpus of tweets that contained just one or two words, with character counts far shorter than Twitter's limit. This would suggest that the pressure to reduce character counts to comply with the message buffer is not as high as expected. That stated, examples of syntactic and morphological variation were found across the corpus, even in short tweets that were at no risk of exceeding the limit. This points to a culture among Twitter users of contracting words and sentences for expressive purposes and for minimising typing effort. This culture effects how words and phrases are reproduced on screen in a number of different ways.

Syntactic and morphological variations in the ILB were conspicuous by their absence. Studies into other web genres such as MUDs (Cherny, 1999), chatrooms (Paolillo, 2001), and micro-blogging sites (Crystal, 2011), commonly find syntactic and morphological variations. However, the language in the ILB – particularly in posts – was for the most part written in fully formed, grammatically correct sentences, albeit with frequent dialectical variations in vocabulary and grammar, and with unintentional grammatical and spelling mistakes depending on the author's standard of Irish. In fact, there were many examples where ILB members had written comments correcting spelling or grammatical mistakes in their own writing or that of others. Far from being seen as a rebuke, these corrections appear to be welcomed by the error-makers, as seen in this example where one blogger (NI) pointed out a mistake in another blogger's (GFG) post title, "A Fhrankie dhíl":

Blog comment 1
NI: dil [guta gairid] Measaim go bhfuil an focal sách tábhachtacht sa chomhthéacs seo le haird a tharraing ar an litriú.
dil [short vowel] I feel the word is quite important in this context to draw attention to the spelling.

Blog comment 2
GFG: go raibh maith agat, ceartaithe agam
thank you, I've corrected it

NI goes on to amend a spelling mistake he made in the initial correction:

Blog comment 3
NI: Maidir le litriú: *aird a tharraingT* (own italics)
Regarding spelling: aird a tharraingT

This practice of correcting grammatical and spelling errors was relatively common, with amendments often introduced by the Irish or Latin words for correction: *coigeartú* or *recte* respectively. It would suggest that members at the core of the ILB fulfil a collective role in maintaining high standards of Irish in each other's blogs.

The practice of correcting spelling or grammatical mistakes was not evident in the Twittersphere. This might reflect the general acceptance of non-standard writing among users in the context of a strict character limit and an ethos of linguistic inventiveness. In the *GA* group, members regularly corrected their own errors, often by posting messages immediately after the offending message clarifying how it should have read, as in the example below:

Facebook message 1
ToM: Dia at [sic] sabhail, an as an cabhán thú Oisin?
God save us, are you from cavan Oisin?

Facebook message 2
ToM: Ar sabhail, brón orm
Save us, sorry

Interestingly the above *GA* member did not correct the lack of diacritic markers or capitalisation in his message, suggesting an acceptance among users that non-standard spellings or punctuations can be used as long as they don't interfere with the legibility or comprehension of the message. The threshold at which messages become incomprehensible is often pushed to the limit, especially in the Twittersphere where, for example, phrases may appear as #hashtags with no spaces between words.

Tweet 1
@username1[88] @username2 aodháin #níltúiagusníraibhtúigcolaistenatrionóidesodúndochlabmarisamadanmórsa lachthú
@username1 @username2 aidan #youarenotinandyouwerenotintrinitycollegesoshutyourmouthbecauseyoua reabigdirtyfool[89]

The above example is an extreme one, however it was common in the ILT and *GA* to omit punctuation as a simple way of reducing character count and typing time and effort. For example, it was common for full stops to be omitted from the end of the last sentences of Facebook messages in *GA* (as in Facebook message 2) and tweets in the ILT (as in Tweet 2), a practice found in English language tweets in Crystal's 2011 study.

Tweet 2
@username1 @username2 @username3 Bhí mé ag ceapadh sin. Sílim gurbh sin - & nathanna ilbhéarlacha – an chúis go bhfuil mo % mar atá
@username1 @username2 @username3 I was thinking that. I think that - & multilingual expressions – are the reason my % is the way it is

In other examples from the ILT, full stops were omitted from sentences at the start and middle of tweets, and capital letters alone indicated the beginning of new sentences. *GA* members also frequently omitted full stops in the middle of messages,

[88] All @usernames in sample tweets have been anonymised.
[89] "You *are* not in and you *were* not in Trinity College so shut your mouth because you are a big dirty fool."

some users expressing themselves in an almost stream-of-consciousness flow of words without pause (as in Facebook message 4).

> Facebook message 4
> BrB: Go deas tá [name] ar dòigh.[90] Obair iontach dèanta aici [name] [name] do fhèile maith iad an chead uair eile a bheadh tù in RnaF Casfar muid ar a cheile maith thù
> *Nice [name] is great. Fantastic work done by her [name] [name] for a festival well done to them the next time you are in RnaF[91] We will meet up well done*

> Tweet 3
> @username1 2.5uair ar bhus! Agus cén fáth go bhfuil NI lódáilte meastú?
> *@username1 2.5hours on a bus! And why is NI (Northern Ireland) loaded do you think?*

In some tweets, spaces were omitted between words (see the words "meas tú" in Tweet 3), particularly where words were preceded or succeeded by numbers (see "2.5uair" in Tweet 3). It is impossible to tell whether such omissions are as a result of sloppy or rushed typing, but their frequency would suggest that they were made deliberately. The Irish language uses one diacritical marker in the form of an acute accent to indicate that vowels are long: á, é, í, ó, ú. This accent is important in differentiating between words that might be spelt the same but pronounced very differently. In general ILT and *GA* members included the accents where required, although not all users included it in every instance and many tweets and Facebook messages have examples of both inclusion and omission (as in Facebook message 2 above). Tweet 4 below is an example of where the accent has been completely dropped, as seen in the words "suíomh", "bhí" and "taispeáint".

[90] It is unsure why this user commonly used the grave accent (more commonly associated with Scots Gaelic) instead of the standard acute accent. Presumably this had to do with the user's keyboard set up and ease of use.
[91] 'RnaF' was a common abbreviation in the *GA* Facebook group for the Donegal Gaeltacht village of *Rann na Feirste*.

Tweet 4

@username1 an suiomh a bhi ar taispeaint ag Antaine an 7ain seo caite. GRMA

@username1 the site that Antaine was showing last week. THX

Some common logograms used as shorthand in other text genres were adopted in the ILT and less frequently in *GA*. The ampersand (&) was used widely as a conjunction, including at the beginning of sentences; and the percentage sign (%) was used sometimes as a standalone word (both features can be seen in Tweet 3). Mathematical logograms such as + and = were used in various circumstances to add or equate tweet content (as in Facebook message 5); and the back slash (/) was frequently used instead of the conjunction "or". This was an effective way of reducing character count without much additional effort required on behalf of the reader.

Facebook message 5

LuM: ní hionann Tír Chonaill agus Co. Dhún na nGall. Co. Dhún na nGall = Tír Chonaill + Inis Eoghain.

Tyrconnell and Co. Donegal are not the same. Co. Donegal = Tyrconnell + Inishowen.

A common practice across the three groups was the use of numerical digits in place of their written counterparts. As in Tweet 3 above, digits can easily be substituted for numbers in their written form, reducing character count and typing effort in the process. The Irish language has three different ways of referring to numbers that change the choice/spelling of words depending on whether the number refers directly to the digit, or is used to count objects or people. Using digits enables web users to avoid having to choose and spell the numbers' written equivalents. This is an example of a linguistic feature common to the ILT that makes it potentially easier for less competent Irish speakers to use the language more effectively in a text-based environment. It should be noted, however, that though the use of digits helps avoid choosing between different number variations, the often complicated rules of Irish grammar regarding lenition and mutation of nouns following numbers were generally adhered to after digits.

Some ILT users (and fewer *GA* members) took the use of digits a step further and

substituted them for syllables in words that sounded similarly. This is a feature common in English language text messaging and CMC. Common examples include "b4" instead of the word "before" and "gr8" for "great". This can help reduce the space and effort needed to communicate on screen. Many numbers in Irish double up as common words when written in their non-digit form: the Irish word "aon" (one) can also mean "any"; "dó" (two) is also a common prepositional pronoun meaning "for/to him"; "trí" (three) can mean "through"; "sé" (six) can mean the pronoun "him" or "it"; and the Irish word "ocht" (eight) is a common sound in many everyday words such as "anocht" (tonight), "bocht" (poor), "locht" (fault), etc. Furthermore, some common words may be substituted by letters that sound similar when read. For example, the verb "bhí" (was) was substituted in the ILT by the letter "V" because it is pronounced the same. Despite the potential for substituting words and syllables for digits and letters in Irish, it was not widely practiced among core members of the ILT or *GA*. Only six of the 16 individuals in the ILT, and one core member of GA, had used this feature and in very few instances. In Tweet 4 above, the first part of the word "seachtain" (week) is substituted for the number seven (spelt "seacht" in Irish) in a hybrid word that appears as "7ain". Similarly, in Facebook message 6 the second syllable in the word "anocht" (tonight) is substituted for the number eight (spelt "ocht" in Irish). In Tweet 5 the word "ar" (on) is substituted by the letter "r". Anecdotally, the hash sign "#" has been used in digital Irish language communication to represent the word "thaiscí" (darling) because the Irish word is pronounced similarly to the alternative English name for the symbol: "hash key". However, this was not found to be the case in any of the groups under study. On the eight occasions where the word "thaiscí" (alternatively spelt "thaiscidh") appeared in the corpus of Irish language CMC it was consistently spelt in full. It appears that despite the popularity of substituting digits and letters for words in English language text messaging, and despite its potential for minimising typing space and effort, it was not common in the repertoire of features used to write CMC in Irish. This is surprising given the popularity of other contracting practices.

Facebook message 6

OiM: bhí muidne chun an bóthar a bhualadh síos go GD an8 ach tá [name] an tinn, bainfidh mé triall as do chlár raidió ina áit mar sin :-)
we were going to head down the read to Gaoth Dobhair tonight but [name] is sick, I'll try your radio show instead then :-)

Tweet 5

@username1 @username2 is féidir a bheith beagnach cinnte nach wil siad ina choinne,ach darliom b'fhearr díriú r phobail a wil gaeilge acu cheana
@username1 @username2 you can be almost certain that they are not against it, but in my opinion it is better to focus on communities that have Irish already

Another feature common to Twitter and Facebook was the ellipsis of the subject and auxiliary verb at the beginning of sentences. Crystal found the ellipsis of subject pronoun and auxiliary verb to be a common feature in English language tweets, particularly in sentences that followed the opening one (Crystal, 2011: 46). This was also frequently practiced in the ILT and *GA*, often by dropping the words "tá mé" (I am) or "tá sé" (it is) from the beginning of sentences. In Tweet 6 the verb and pronoun have been omitted; a fully formed version would read: "Beidh mé i mBÁC ar an Luan, agus ag pilleadh ar Ghaillimh ar an Mháirt". In this instance, omission of the subject pronoun and its auxiliary verb and the conjunction in the middle of the sentence has shortened the tweet by 14 characters. In Facebook message 7, again a "tá sé" has been omitted from the beginning of the sentence. The overall effect of these practices is to produce the type of short snappy sentences one might expect in spontaneous speech. This adds to the immediacy of the genres, giving the impression that the users are commenting on events as they occur contemporaneously around them. Whereas comment in traditional written media, online newspaper articles and (some) blogs are generally written in the past tense offering reasoned and judged analysis of events that have passed, Twitter and Facebook are full of musings and half-conceived thoughts as they occur to users. This type of clipped message seems to encourage quick and casual response, as happens in the threads of conversations that can build up almost synchronously on Twitter and Facebook in a process likened to chitchat. However, the immediacy of many such messages also prematurely ages

them. Blog posts in contrast are frequently read and commented on days or weeks after they have been published.

Tweet 6
i mBÁC ar an Luan, ag pilleadh ar Ghaillimh ar an Mháirt.
[I will be] in Dublin on Monday, [and] returning to Galway on Tuesday.

Facebook message 7
FrC: Thar a bheith níos fearr ná an bodhrán..goa
Much better than the bodhrán..lol

The morphological and syntactic variations above were most commonly found in the ILT and *GA*. This may be as a result of the more informal chitchat taking place on Twitter accounts and Facebook walls, in contrast to the more formally structured written discourse composed in blog posts. Comments on blog posts were usually more conversational in style and more likely to share the non-standard features identified in tweets and Facebook messages. Where syntactic and morphological variations did occur in the ILB in common with the other two genres, they were most often found in the non-standard spelling of words to mimic the users' own speech. Words were often abbreviated, or letters or syllables were omitted, reflecting the authors' own accents or dialects. These changes most often occurred in common everyday phrases as in the following examples from four different authors in the ILB:

1) 'bhFuil tuilleadh eolais faoin gceol ag éinne?
2) Maith dom m'aineolas, ach cá'l sin?
3) maith agat arís, a Áine!
4) …tá's 'ainn cé a bheidh an príomhpháirtí…

The full written versions of these phrases would read:

1) **An bhfuil** tuilleadh eolais faoin gceol ag éinne?
2) Maith dom m'aineolas, ach **cá bhfuil** sin?
3) **Go raibh** maith agat arís, a Áine!
4) …**tá a fhios againn** cé a bheidh an príomhpháirtí…

Across the three genres it was common to insert apostrophes to indicate where words or letters had been dropped. This reduction of words to contracted forms in formulaic phrases was also documented by Cherny (1999) in her study of English language MUDs. It would appear that where ILB members feel comfortable playing with syntactic and morphological variation it is most often in everyday phrases where the meaning is unlikely to be misinterpreted. Examples abound in the ILT. In Tweet 5 above the word "bhfuil" (is) is written "wil" as it is pronounced phonetically by the speaker, despite there being no letter "w" in the Irish language. Elsewhere in the ILT the same word has been spelt "vuil", this time by a user who pronounces the word with a hard "v" sound according to their own dialect.

> Tweet 7
> @username1 ag obair luath ar maidin. má thosaím orthú, ní bheidh stop orm! céard atá ar bun agat?? g'iarraidh tae??
> *@username1 working early tomorrow. if I start on them, there will be no stopping me! what are you doing?? want tea??*

In Tweet 7 the phrase "Do you want tea?" which in standard grammatical form would read "An bhfuil tú ag iarraidh tae?" has been contracted to "g'iarraidh tae?" Not only has the subject pronoun and verb been omitted, but the words "ag" and "iarraidh" have been combined into one word, again mimicking how the question might sound in informal speech. The same user has added an acute accent to the "u" at the end of "orthu" (on them) in a non-standard spelling that gives a clue as to the speaker's spoken dialect. Many non-standard terms and contracted forms in the *GA* Facebook group also gave clues to the authors' spoken dialects. Because core members of the group were predominantly from Ulster, non-standard spelling and vocabulary most often displayed features of the Donegal Irish dialect. In Facebook message 8, as well as dropping diacritical markers, punctuation and capitalisation, the author has devised her own spelling of the informal phrase "s'againne", a term used commonly in Donegal Irish to mean 'ours'. Moreover, the words "amharc" (look) and "galánta" (beautiful) are common words in the Donegal dialect.

Facebook message 8:

BrB: Maire sagaine ag amharc galánta realt ga

Our Máire looking beautiful, the star of GA.

2. ABBREVIATIONS & ACRONYMS

There were a number of types of acronyms and abbreviations that were shared across the three groups. These can be summarised as follows:

- Acronyms were commonly used to express emotions or everyday phrases, much like in English language CMC.
- Acronyms were commonly used for organisations, in English or Irish depending on the author's preference.
- Irish language acronyms were commonly used for placenames, referring to countries, cities and villages depending on the familiarity of group members.
- English language acronyms were most commonly used for terms related to technology or science.
- Other abbreviations common to traditional written genres were used. However, there were no standard forms for these abbreviations in the online groups, and their spelling varied between genre and users.

Trends in the use of acronyms and abbreviations will now be described in greater detail with examples.

Tweet 8

@username1 Ooh, ba bhreá liom (ag braith ar am, data 7rl dár ndóigh). Seol rphost chugam ag [e-mail address] ldt. :)

@username1 Ooh, I'd love to (depending on time, date etc of course). Send an email to me at [e-mail address] pls. :)

Across the three groups, both acronyms and abbreviations were commonly used to minimise typing time, space and effort. Tweet 8 includes examples of three different types of abbreviations and acronyms in the words "7rl" (etcetera), "rphost" (email) and "ldt" (please), which in standard written form would appear as "agus ar uile", "ríomhphost" and "le do thoil" respectively. The benefits to the author are obvious,

with a total reduction in typing of 21 characters. The acronym "ldt" above is an example of the shorthand used in CMC and text messaging for everyday phrases or emotions. These are a common feature of Crystal's (2006) Netspeak, and can be likened to the use of acronyms such as LOL and BTW in English language CMC. In fact many such acronyms found in Irish language CMC were direct translations of their English equivalents. They appeared in all three groups, written in both capital and lowercase letters. In the ILB this type of acronym was most likely to be found in comments rather than posts, reflecting the more immediate reactions people give in their (short) comments compared to the more considered content of (longer) posts. Many such acronyms were unique to Irish language CMC and are unrecognisable or difficult to decipher by unfamiliar readers. The most common acronym across the three groups was:

GRMA: go raibh maith agat / *thank you*

This was often supplemented with additional letters to intensify the expression, as in GRMMA/GRMMMA: go raibh míle (míle) maith agat / *literally, a thousand (thousand) thank yous.*

The second most common instance of acronyms of this type was used to express laughter:

ASG: ag scairteadh gáire / *bursting out laughing (i.e. LOL in English)*
AGOA/GOA: (ag) gáire os ard / *laughing out loud*

Here, there was a difference between the ILB and the other two groups, with the acronym ASG favoured by bloggers, while (A)GOA was used more frequently by Twitter and Facebook users. Alternative versions included MMFAG: mo mharú féin ag gáire / *killing myself laughing*, and ILG: i lagracha gáirí / *weak laughing*. It is unlikely that these latter acronyms would be readily deciphered by readers, particularly newbies, and they appeared very infrequently.

Other common acronyms included:

DAS: dála an scéil / *by the way (i.e. BTW in English)*
CGL: ceart go leor / *OK*
LDT/LBD: le do thoil/le bhur dtoil / *please (to an invididual and group respectively)*

ILT users regularly peppered their tweets with less common acronyms, confirming the ethos of linguistic inventiveness on the site. References to God appeared in no fewer than four different acronyms – LCD: "le cúnamh Dé" (with the help of God), BLD: "buíochas le Dia" (thanks be to God), OMD: "Ó mo Dhia" (a non-standard gaelicisation of the common English interjection 'Oh my God') and the cruder OMFD: "Ó mo f**king Dhia". Single instances of the following acronyms were also noted - tgbé: "tóg go bog é" (take it easy), and psé: "pé scéal é" (anyway). In using acronyms like these, the author assumes a mutually understood register peculiar to online communication and shared with other members of their group.

Not all acronyms were unique to CMC. Many acronyms were used across the three groups as they would appear in traditional written genres, particularly the acronyms of organisations. These were written in English (e.g. BBC: British Broadcasting Corporation, IFTA: Irish Film and Television Award, NCT: National Car Test) or in Irish (e.g. SF: Sinn Féin, TD: Teachta Dála, RTÉ: Raidió Teilifís Éireann) according to how they most commonly appeared in other genres. Other acronyms were used variously in English or Irish depending on the preferences of the authors. For example, University College Dublin was variously referred to as UCD or COBÁC, the Gaelic Athletic Association appeared as both GAA and CLG, Northern Ireland was either NI or TÉ, and even Young Fine Gael appeared as YFG or FGóg ("óg" meaning "young" in Irish). This reflects the general flexibility of users in the three groups to recognise both English and Irish acronyms. It was common to use acronyms as shorthand for frequently referred to places. One particularly common abbreviation in the ILT was of the Irish name for Dublin – Baile Átha Cliath – which appeared as BÁC in many tweets. It was common to lenite or mutate the acronym in Irish tweets according to the rules of Irish grammar, for example, "i mBÁC" for "in Dublin", "ó BhÁC" for "from Dublin" and the non-standard "muintir BHÁC" for "the people *of* Dublin". In the *GA* Facebook group, the acronyms GD and RnaF were used

commonly to refer to two Gaeltacht areas of Co. Donegal - Gaoth Dobhair and Rann na Feirste respectively. In the ILB, the acronyms SAM (equivalent to English language "USA") and AE (equivalent to English language "EU") were commonly used when referring to news or politics in those regions.

One area where English language acronyms prevailed was around the theme of technology and science. This reflects the English-speaking world's dominance in the development and spread of new computer hardware and software, and the popularity of the English language in publishing and promoting new scientific research. English acronyms were used to describe computer software like API, UTF-8, ASCII; web functions like RSS (rich site summary) and FB (Facebook); general computer terms like IT (Information Technology), PDF (portable document format) and CD (compact disc); and scientific terms like GM (genetically modified) and DNA (deoxyribonucleic acid), without any effort to compose Irish language alternatives. The Twitter-specific acronyms DM (direct message), RT (retweet) and FF (Follow Friday, used to promote other users on Twitter) almost always appeared in their English acronymised forms, even when referred to in the middle of tweets when not carrying out their Twitter functions. It remains to be seen whether Irish language equivalents will evolve over time to denote these common activities in Twitter.

Some popular abbreviations across the three groups were common to other written genres, such as the abbreviation "lch" for "leathanach" (similar to "p." for "page"), "srl" for "agus ar uile" (similar to "etc." for "etcetera"), "m.s." for "mar shampla" (similar to "e.g." for "for example"), "pic" for "pictiúr" (similar to "pic" for "picture"), and "r-phost" for "ríomhphost" (similar to "e-mail" for "electronic mail"). These examples give abbreviated forms of words or phrases that the reader is used to seeing abbreviated elsewhere, in Irish and English writing. These common abbreviations had no standard spelling across the online groups as they do in traditionally edited writing, however, but rather were left to the interpretation of the authors. The abbreviation "m.s." ("e.g." in English) appeared alternatively as "ms", "msh" or "m.sh." in the ILT. Similarly, "srl" ("etc." in English) was written variously as "srl", "&rl" or "7rl" across the three groups. Other less common abbreviations appeared where they were deemed by the author to be clear from the context. The word "bliana" (years) appeared as "bl", "deireadh seachtaine" (weekend) as "ds", and "teachtaireacht" (message) as "tcht". The use of such abbreviations could act as a help or hindrance to less competent Irish speakers. On the one hand readers must be able to

decipher the abbreviations as they come across them; on the other hand, when composing messages themselves they may make use of such abbreviations to avoid difficult spelling or grammatical errors.

Finally, members of the *Gaeilge Amháin* Facebook group frequently referred to the group as 'GA' in their messages. When this was written in lowercase letters it was sometimes difficult to distinguish, especially when the author dropped other punctuation, as in Facebook message 9.

> Facebook message 9
> BrB: Bìgì linn i gconai ar ga bheadh muid cailte gan bhuir cuideachta bheadh Ga ro uaighneach gan sibh agus a chaidre ga ar fad
> *Stay with us on GA we would be lost without your company GA would be too lonely without you and all of the GA friends*

GA members also referred to themselves frequently using abbreviated forms, so one member named Bridanna was referred to in a message as 'BA' and another called Bríd Eilís was referred to as 'BE'. As well as cutting down on typing space and effort, the use of such abbreviations implies a familiarity between users.

3. PLAY WITH MODALITY

There were a number of ways in which members of each group played with the modality of their messages. These can be summarised as follows:

- Emoticons were frequent in all three groups (in the ILB, more often in comments than posts).
- Multiple exclamation and question marks were commonly used to heighten the intended emotion of the message.
- Capitalisation was commonly used to add stress to specific words.
- Other graphic devices including italicisation and asterisks were used to single out or add emphasis to words.
- Ellipsis points were used to add pauses during sentences.
- Interjections were frequently included in messages. These were sometimes taken directly from English examples; while in other instances they were

adapted fully or partly to the rules of Irish phonetics.

- Additional letters were added to words to mimic syllable stress in speech.

Various features of play with modality from the three groups will now be described in greater detail with examples.

Play with modality describes the written features that denote the mode, mood, or manner in which posts and comments are made. In the absence of audiovisual cues, web users use innovative graphical variations to reflect the emotion, intonation, stress, speed, rhythm, pause, and tone of voice that occur in face-to-face communication (Crystal, 2006). These features go some way in preventing readers from misinterpreting the tone of messages from how the writers intended. One feature common to all three groups was the use of emoticons: symbols intended to represent various facial expressions. The emoticons in Irish language CMC are not unique, but use symbols for smiling and frowning faces in the common Western style of English language CMC, with colons and brackets/letters/digits representing a smiling or frowning face when viewed from the side. The following examples show how emoticons have been used to express humour, disappointment and sympathy respectively:

Facebook message 10
OiM: sin an áit ar pháirceáil mé é! :-P
that's the place I parked it! :-P

Blog comment 4
NI: Tá sé in ainm a bheith greannmhar :-\
It is supposed to be funny :-

Tweet 9
sil: @username1 ouch! :(na héisc bhochta!
@username1 ouch! :(the poor fish!

There were no rules governing what emoticons should or should not be used, and individual users adapted various characters and letters to their own ends. In the ILT, for example, some were used to denote nervousness through the characters :-s and

bewilderment through the characters o.O - the latter meant to symbolise a raised eyebrow.

Other graphical features were used with words to express different intonations and stresses. Again, these features are common to English language CMC. They included the use of multiple exclamation marks and/or question marks to denote emphasis, excitement, sarcasm, or puzzlement as in the following examples:

Blog comment 5
acm: An é fíor a rá go bhfuil an pointe is faide ó thuaidh in Éirinn sa Deisceart!!!!!!!!?????
Is it true to say the most northerly point in Ireland is in the South!!!!!!!!?????

Facebook message 11
AnK: Na Gardaí Síochána : Amadáin !!
!!
!!!!!!!!!!!!!!!!!!!!!!!!!!
The Police Force : Fools !!!
!!
!!!!!!!!!

Tweet 10
Anocht @username1 SRAITH NUA 'Rásaí Na Gaillimhe 2'!!! Ag tosnú anocht ag 9.30pm!!! Greanmhar, cliste agus Iontach!!!! Ná cailigí ea!! #tg4
Tonight @username1 NEW SERIES 'The Galway Races 2'!!! Starting tonight at 9.30pm!!! Funny, clever and Brilliant!!!! Don't miss it!! #tg4 [Irish language TV channel]

In general exclamation marks were far more common across the three genres than in traditional written genres. For example, in a corpus of almost 2,500 tweets there were over a thousand exclamation marks, with one tweet using six in a row. In a corpus of 2,975 Facebook messages there were 1,870 individual exclamation marks, with one message (Facebook message 11 above) ending in 156 exclamation marks in a row.

Although these marks were used to heighten the sense of emotion, their overuse in many instances acted to diminish their power, with some messages using exclamation marks in seemingly banal everyday sentences. Tweet 10 above is an example of exclamation mark overkill, with multiple exclamation marks used repeatedly in a tweet that is essentially providing information about the start of a new television series.

Capital letters were commonly used to stress specific words in sentences. These were often used to mimic the sound of a raised voice, as in Blog comment 6 and Facebook message 12.

> Blog comment 6
> CHD: Ach rud beag amháin – tá's agam go ndeirtear fúinn go mbímid róghoilliúnach anseo i gCorcaigh, ach fós féin NÍ HIONANN STOUT AGUS GUINNESS. Tá Murphys agus Beamish ann leis.
> *But one small thing – I know that they say we are too sensitive here in Cork, but still STOUT DOES NOT EQUATE TO GUINNESS. There's Murphys [sic] and Beamish too.*

> Facebook message 12
> MaB: Bí maith, bí ciallmhar ach ná bí …… STUAMA ! Bain sult as a chara !
> *Be good, be sensible but don't be ..... LEVELHEADED ! Enjoy it my friend !*

In other instances word stresses were denoted by marking the word with asterisks. This feature was also used to denote written actions or sound effects. Again this is a feature common to English language CMC and is used in two distinct ways. Firstly, as in Tweet 11, asterisks are placed on either side of a word to add emphasis. In this instance the author is not entirely sure about what they are tweeting and singles out the word "dóigh" (think) to highlight this uncertainty. Thus asterisks can be used in instances where words might be italicised, underlined or written in bold in traditional written genres. The second common use of asterisks is to indicate where an action is being demonstrated. Blog comment 7 has such an example where the action of sighing is highlighted with asterisks. In other examples from the ILT the act of

exhaling (*asanalú*) and grabbing something (*yoink*) were treated in this way.

Tweet 11
@username1 @username2 ní *dóigh* liom go bhfuil aon cheo mar sin aige go fóill, sílim go raibh sé i gComhar cheana, ach an-seans go mbeid
*@username2 @username2 I don't *think* he has anything like that yet, I think it was in [the magazine] Comhar before, but every chance it will be*

Blog comment 7
NI: Ní fheicim mo bhlagsa […] ar an liosta ar dheis. *osna*
*I don't see my blog […] on the list on the right. *sigh**

Another common orthographical feature was the frequent use of ellipsis dots, also noted in Crystal's (2011) study of English language tweets. In traditional written genres this would commonly indicate the intentional omission of some text. However, in CMC ellipsis dots are used more flexibly, often to mimic the flow and intonation of speech in the text. It is common for ellipsis dots to be used mid-sentence, for example, to break up the rhythm of the writing. This may indicate a pause for thought on behalf of the author (as in Facebook message 13), or build suspense before a revelation however big or small. In Tweet 12, nine ellipsis dots are used to indicate some hesitation on behalf of the user to continue with what she assumes to be a stupid question. When they appear at the end of a tweet (as in Blog comment 8 below) the function appears to be one of finishing a thought without any real closure, the sense of someone trailing off before they've reached a satisfactory conclusion. This may act to encourage others to join in the conversation and offer their own thoughts on the topic.

Facebook message 13
AnK: 1972… ni raibh me ach…? bliain d'aois lol!!!!
1972… I was only…? one year old lol!!!!

Tweet 12
@username1 ok brón orm…… ach cé hé Tom Brady? ooppss!
@username1 ok sorry…… but who is Tom Brady? ooppss!

Blog comment 8

RAM: Tugtar an 'grandmother hypothesis'[92] air. Hmmm…

They call it the 'grandmother hypothesis'. Hmmm…

It was very common to include interjections in tweets, expressing a wide range of reactions from surprise and pleasure, to sympathy, sighs and gasps of pain. We've already seen in Tweets 8, 9 & 12 and Blog comment 8 examples of English language interjections (Ooh, ouch, ooppss, Hmmm), spelt with combinations of letters that don't follow the rules of Irish spelling and dropped into otherwise Irish language content. It is clear from these and other examples, that a lot of users in many cases feel more comfortable expressing the sound effects and reactions associated with speech through common English language interjections. These may be the types of interjections users have witnessed in English language CMC elsewhere online. That said, there are many examples of Irish language interjections in the three groups, or at least interjections spelt using Irish phonetics, albeit sometimes with the use of letters alien to the Irish alphabet (i.e. j, k, q, w, y, z).

Blog comment 9

CHD: Iom-iom-IOM! Tá uisce lem' fhiacla is mé ag breathnú ar an bhfíseán san.

Yum-yum-YUM! My mouth is watering looking at that video.

Facebook message 14

AiC: yéé háá , cad chuige nach bhfuil hata air?

yee haa , why does he not have a hat on?

Tweet 13

squíí :D tá mo mhamaí ag teacht go BÁC ag an d.s., muid ag stopadh in óstán agus tá siad ag cur bricfeasta vegan ar fáil domh :D

squee :D my mammy is coming to Dublin on the w.e., we are stopping in a hotel and they are going to make me a vegan breakfast :D

[92] The underline indicates that the words appeared as a hypertext link.

The above three messages have examples of gaelicised interjections. In a creative use of Irish spelling in Blog comment 9, the English interjection 'yum' has been spelt 'iom' according to the sounds of Irish. The third 'iom' as been capitalised to intensify the author's appreciation for the food being discussed. In Facebook message 13 the author has created a quasi-Irish version of the English expression 'yeehaw' using diacritic markers that make sense in Irish, while keeping the non-Irish letter 'y'. Tweet 13 takes the English interjection "squee", commonly used in CMC to express reaction to something cute on screen, and partially gaelicises it by changing the double-e to a double-í, again conforming to the rules of Irish phonetics. However, the letter "q" – alien to the Irish alphabet – is kept, adhering to the general shape of the interjection as it appears in English CMC. Using a combination of lowercase and capital letters, and letters in (á, í) and out (y, q) of the Irish alphabet the above messages are creative interpretations of the sounds of speech, but spelt in ways that make sense specifically in the context of Irish language phonetics. It suggests a desire among some members of each group to adapt their writing to be monolingually Irish, while maintaining the freedom to express gasps, cries, sighs and emotions not usually seen in Irish language writing. It shows a willingness and confidence among Irish web users to engage in new and innovative ways of expressing the sounds and intonations of speech in the confines of text-based on-screen interactions.

Facebook message 15
MaB: MISE !!!! ÁÁÁÁÁÁÁÁÁÁÁÁÁÁÁÁÁÁ
ME !!!! AWWWWWWWWWWWWWWW

Blog comment 10
RAM: Ar ndóigh, níor thuig mé a dhath de, ach bhí an bua acu – wúhúúúúúúúúúúúúúuu!
Of course I didn't understand any of it, but they won – woohooooooooooooooooo!

Tweet 14
@username1 ugggggggggh is fuuuuuuuuuath liom sin!!
@username1 ugggggggggh I haaaaaaaaate that!!

Another common feature was the use of additional letters to extend certain syllables in key words. This acted not just to single out words for emphasis, but to mimic how the author might stress the extended syllable in speech. This practice was particularly common in interjections where letters were added to amplify the sound and the emotion imbued in the words. This is the case in Facebook message 15 and Blog comment 10 above, where the gaelicised versions of the English interjections 'aww' and 'woohoo' have been given greater stress with the repetition of the letters 'á' and 'ú' respectively. Two dramatic instances can also be seen in Tweet 14 where both the English interjection "ugh" and the Irish word "fuath" (hate) have been given an additional eight g's and u's respectively. The emphasis of the word "fuath" in particular leaves us in no doubt of the individual's dislike for the topic being discussed.

4. LANGUAGE ROUTINES

The researcher found examples of language routines unique to each group. These can be summarised as follows:

- ILB members supported each other through short formulaic messages offering words of welcome, and expressions of humour, sympathy and praise.
- In the ILT, routines have emerged around the specific use of hashtags and @usernames in Twitter. These have been adapted to the Irish language in different ways.
- In the *GA* Facebook group, routines have emerged around both welcoming new members and re-inforcing the language rules associated with the group in different ways.

These language routines from the three groups will now be described in greater detail with examples.

Language routines are standardised patterns of speech that are frequently uttered in a speech community. They act as group rituals that give a feeling of membership and a sense of confidence to a community, as well as behavioural certainty to their members (Cherny, 1999). Routines emerged in response to posts in the ILB broadly

around four themes, where commenters wrote short expressions of: *welcome* – to new or returning bloggers; *praise/encouragement* – e.g. when bloggers posted photographs they had taken or poetry/prose they had written; *amusement* – e.g. at a blogger's jokes or anecdotes; or *sympathy* – e.g. when bloggers announced an illness or death. An example of a routine can be seen in the response to one blogger's post about the death of a parent:

> Blog comment 11
> CLR: Mo chomhbhrón ó chroí leat a chara.
> *My sympathies from the heart, my friend.*
>
> Blog comment 12
> GFG: Ní maith liom bhur dtrioblóid.
> *I'm sorry for your troubles.*
>
> Blog comment 13
> CHD: Suaimhneas síoraí dá hanam.
> *Eternal peace for her soul.*
>
> Blog comment 14
> HNY: Ní maith liom do bhris.
> *I'm sorry for your loss.*

And the simple response from the blogger:

> Blog comment 15
> RNG: Go raibh maith agaibh, a chairde.
> *Thank you, friends.*

In contrast to comment threads that engage members in conversations around the themes of posts, routines comprise of short formulaic messages, often a common expression, and sometimes just a single abbreviation. They are not intended to elicit further comment, but simply act to acknowledge the blogger's presence, achievement, humour or struggle. In an example of a routine response to humour, one blogger

received four expressions of the acronym ‘ASG’ and one ‘gáire os ard’ (both versions of the English language expression ‘laugh out loud’) in reaction to a joke he posted. These formulaic responses were sometimes followed by a line or two of further comment. In another example, a blogger received five short comments wishing her well on the event of her birthday. The blogger herself then responded to each comment with a short message of thanks. In this way, language routines bond members through a shared sense of involvement, support and solidarity, without commenters having to invest large amounts of time or effort in forming unique and original responses or engaging other users in dialogue.

In the ILT, as tweeting has become increasingly popular and more widely used, particular routines associated with the use of hashtags and the @username identifier have emerged. Hashtags have evolved in Twitter to perform a range of functions. These have been adopted in the ILT and adapted in some cases to the rules of Irish grammar.

> Tweet 15
> tá 7LÁ an-mhaith anocht. Anois ar TG4 #gaeilge
> *[the current affairs TV programme] 7LÁ is very good tonight. Now on TG4 #irish*

> Tweet 16
> #charliebird #rte #gaeilge #teip Charlie Bird ag caint ar Radio 1 faoi na meáin nua, Gan ionadaí ó na meáin Ghaeilge. #anghnathrud
> *#charliebird #rte #gaeilge #fail Charlie Bird talking on Radio 1 about the new media, Without a representative from the Irish language media. #theusualthing*

Hashtags can be used as a quick and easy way of searching for tweets that share a specific topic or theme. By creating a hashtag that categorises the subject matter of the tweet, users can promote their visibility in the Twittersphere among other users who might share similar interests. This process is described by Ruth Page in her analysis of narrative strategies in English language tweets as ‘folksonomic tagging’ (Page, 2012). One common example from the ILT that falls into this category is the

hashtag "#gaeilge" where the Irish name for the language itself is used to tag tweets, as seen at the end of Tweet 15. In fact, from the total corpus of almost 2,500 Irish language tweets there were 118 individual uses of the #gaeilge hashtag. This hashtag enables users with an interest in the language to search on Twitter for recent Irish language content, with all of the tweets containing #gaeilge listed in reverse chronological order. Because of the breadth of topics that could fall under the theme #gaeilge, the results could cover a wide range of subject matters in Irish and other languages. However, at their root would be a collection of tweets and users that shared a common interest in the Irish language. More event-specific hashtags are commonly created around television and radio programmes. In Tweet 16, the author used hashtags identifying a media organisation (#rte) and a specific news reporter (#charliebird) to draw attention to his opinion about a debate taking place on radio. Across the corpus of Irish language tweets there were 19 instances of the #tg4 hashtag, commenting on content from the Irish language television channel TG4.

Tweet 17
sílim gur andúileach feamainne mé. #íumm #drabhlás #rocknroll
I think I'm a seaweed addict. #yumm #debauchery #rocknroll

As well as grouping tweets folksonomically, hashtags perform another function in Twitter – that of creative expression. Page describes these as "idiosyncratic examples which function as expressive punctuation" (Page, 2012: 4). In Tweet 16 above, the hashtags "#teip" (fail) and "#anghnathrud" (the usual thing) were not used to navigate readers to the tweet. It is highly unlikely that someone would search for #anghnathrud on Twitter. In fact, a standardised spelling of 'the usual thing' in a hashtag would be '#angnathrud' without the initial 'h'. The misspelling is inconsequential, however, as other users are very unlikely to search for that specific phrase. Rather, these hashtags were used simply to express the author's disappointment and resignation that the radio programme he's commenting on has failed to include an Irish language commentator. In Tweet 17 above, the Irish language hashtags "#íumm" (another Irish language interpretation of the English interjection 'yum') and "#drabhlás" (debauchery) are unlikely to be searched for; rather, they express the author's pleasure in eating seaweed and her so-called 'rock 'n' roll' lifestyle in a tongue-in-cheek fashion.

Tweet 18

grma! :) x "@username1: #ff @username2 @username3 @username4 @username5 @username6 @username7 @username8 @username9 @username10

thx! :) x "@username1: #ff @username2 @username3 @username4 @username5 @username6 @username7 @username8 @username9 @username10

One specific hashtag that serves an entirely different function is the common hashtag #FF, which stands for 'Follow Friday'. This is used widely across the English language Twittersphere as a way of promoting a selection of other Twitter accounts that they deem worthy of interest to their followers (Page, 2012). It was found in ten tweets in the ILT corpus. A Follow Friday tweet is usually posted on Fridays in the form of a list of promoted @usernames preceded by the hashtag #FF or #ff. In Tweet 18 above, the author is thanking another user for including them in their Follow Friday. All but two of the nine users listed in the original Follow Friday tweet are Irish language Twitter users according to Indigenous Tweets. In this way the #FF hashtag can be used within a language community to help introduce different users who share that language.

Both hashtags and @usernames commonly appear mid-sentence in English tweets, as users and topics are referred to naturally within the body of messages. There are consequences to this practice in Irish, however, as decisions have to be made about whether or not to adhere to the rules of lenition, mutation, the genitive case and the vocative case associated with Irish grammar. Since tweets are automatically sent to the @usernames that appear within the tweets, changing the spelling of the username after the @ sign would mean that it would fail to reach its intended target, or potentially be sent to another user entirely. So, for example, if the user "@peadar" was addressed in a tweet as "a @pheadair" according to the rules of the Irish vocative case, the tweet would inadvertently be sent to the user "@pheadair" if such a user existed. Similarly, if the function of a hashtag is the folksonomic one of categorising it according to its subject matter, changing the spelling of the root word after the # symbol would make it far more difficult to search for the tweet. So, for example, if the hashtag "#gaeilge" occurred mid-sentence after the word "sa" meaning "in the" it

ought to be spelt "#ghaeilge" according to the rules of lenition in Irish. However, this would mean the tweet would no longer appear in the results of a search for the hashtag "#gaeilge" as was originally intended. For these reasons the vast majority of @usernames and hashtags are left unlenited, unmutated and unchanged in instances where the genitive and vocative cases would normally apply.

> Tweet 19
> Ní dóigh liom go n-aontóinn le @username1 cén lá don tseachtain atá ann, ach by dad tá scéalta aige i n @gaelsceal
> *I don't think I'd agree with @username what day of the week it is, but by dad he has some stories in [Irish newspaper] @gaelsceal*

> Tweet 20
> @username1 breathnais idir lamha agam anois don eisteddfod, fonn ort teacht? is le h @username2
> *@username1 learning welsh at the moment for the eisteddfod [Welsh festival], do you want to come? and with @username2*

However, three core members of the ILT were found to have made some efforts to adapt @usernames to the rules of Irish grammar. In Tweet 19 the initial consonant mutation "n" was placed before the word "gaelsceal" as it normally does when a word beginning with "g" is preceded by the word "i" (in) in Irish. However, rather than placing the "n" directly in front of the "g" of "gaelsceal", thereby interfering with the spelling of the @username, the "n" was placed in front of the @ sign. Thus, "i ngaelsceal" becomes "i n @gaelsceal": an innovative way of adhering to the rules of Irish grammar without interfering with the specific function of the @username in Twitter. In Tweet 20 above, the word "le" should add a "h" in front of @username2 as it begins with a vowel. Again this was added before the @ sign thereby conforming to grammatical rules, albeit with a non-standard structure. These examples are in the minority, but they demonstrate a commitment among some core ILT members to maintain as monolingually Irish a space as possible albeit using new orthographic innovations.

Language routines in *GA* existed around welcoming new members. This began with an automatically generated message announcing that a named established member had added a named new member to the group. This message appeared as in Facebook message 16, in whatever language the reader had chosen as their default language on Facebook.

Facebook message 16
Chuir [name of established member] [name of new member] leis an ngrúpa.
[Established member] added [new member] to the group.

These were typically followed by other established members – almost always the group administrator and a small number of prolific members – posting formulaic messages of welcome to the newcomer in response. The new member commonly responded to such messages with their own message of gratitude. Such a thread of welcoming messages can be seen in Facebook messages 17-20.

Facebook message 17
AiC: Fáilte isteach chuig GA a chara.
Welcome to GA, friend.

Facebook message 18
MaB: Fearadh na fáilte romhat a chara.
A hearty welcome to you, friend.

Facebook message 19
PeM: Fáilte mhór chuig G
A big welcome to G[A]

Facebook message 20
New member: gur raibh [maith] agaibh :)
Thank you :)

As well as welcoming new members, the administrator regularly posted birthday wishes to some members of the group. These would routinely be responded to with messages of congratulations by other members acquainted with the person celebrating their birthday.

Another routine unique to *GA* is the regular messages posted to the group wall reinforcing the language rules of the group. These were usually posted by the group's administrator – often appearing in the same repeated format – reminding members of their duties to contribute messages in Irish only. Facebook message 21 is typical of this kind of message. In other instances other core members would post their own messages reinforcing the rules and ethos of the group, often explicitly in support of the administrator's efforts.

> Facebook message 21
> AiC: A Chairde GA
> Ba mhaith liom sibh bheith dílis d'ainm an tsuímh seo … Gaeilge Amháin. Ciallaíonn sin go scríobhaimid i nGaeilge agus go gcuirimid suas nascanna atá sa Ghaeilge […]
> *Dear Friends of GA*
> *I would like you to remain true to the name of this site ... Irish Only. That means that we write in Irish and we put up links that are in Irish […]*

This particular message was 'liked' six times by other *GA* members. It was also routine for other members to post responses to such messages, commending the author, agreeing with the principle, arguing why such a rule was necessary, or elaborating on the ethos of the group. The above message from the administrator attracted the responses listed in Facebook messages 21-26. These are typical of the short, often formulaic, messages of support that were posted by various group members.

> Facebook message 21
> GnD: an ghaeilge abú :-)
> victory to the irish language :-) [rough translation]

Facebook message 22
AnK: Ag tíocht lem' thuairmi fhéin a [name] . Maith thu !
Corresponding to my own opinions, [name] . Well done !

Facebook message 23
AlC: mise fostà !
me too !

Facebook message 24
MaB: Grma a [name]. Aontaím go hiomlán leat.
Thx, [name]. I completely agree with you.

Facebook message 25
BrB: Maith thù dia go deo libh
Well done long may you prosper

Facebook message 26
AlC: Béarla = truailliù anseo…
English = pollution here…

The above thread of messages demonstrates the value placed by members of *GA* on the Irish-only ethos of the group. By regularly reminding group members of their duty to communicate solely through Irish and by engaging others in praising the rules associated with it, *GA* remains a monolingually Irish space located within a network of billions of Facebook users communicating with each other in a myriad of more widely spoken languages.

5. CODE SWITCHING

There were many different code switching practices found in each group. These can be summarised as follows:

- Messages posted in the ILB and *GA* are primarily Irish only.
- Members of the ILT are most likely to tweet in both Irish and English, often

mixing both languages in their tweets.

- Members of all three groups occasionally switched to English for some common words and interjections.
- Members of all three groups occasionally switched to English for uncommon, difficult or specialist vocabulary.
- Code switching for words relating to technology or science was particularly common.
- Words in languages other than Irish were often marked apart using different practices.
- Web users often included translations of difficult or specialist words in their messages through different practices.

Code switching practices from the three groups will now be described in greater detail with examples.

The majority of ILT members (63%) tweeted in Irish less than half the time. It is unsurprising, therefore, that switching to English (and other languages) in the middle of tweets was common. However, in the ILB core and *GA* where members were committed to communicating monolingually in Irish, switching occasionally to English was also common. In fact, although the ILB core and *GA* are primarily Irish language spaces, there were many instances recorded where words, expressions, phrases and proper nouns from other languages – ovewhelmingly English, but including other widely spoken, minority and ancient languages – were peppered throughout the content. It is not surprising that English words and phrases appeared in the ILB, given that most ILB members lived in predominantly English-speaking countries. However, the frequency with which other languages appeared was unexpected. This is potentially explained by several factors: many ILB members lived in non-English speaking countries; some members worked or studied in university linguistics departments; language was a common blog topic in the ILB; and, as Irish speakers, some ILB members may have had a natural appreciation for lesser-used languages and/or a resistance to the over-dominance of English. The occasional switching to English words in the *GA* Facebook group occurred despite the explicit rules – regularly endorsed publicly by group members – requesting members to communicate in Irish only.

Code switching was most common in four specific contexts. We have already seen how English language acronyms (associated with Netspeak, e.g. lol; organisations, e.g. BBC; and technology, e.g. PDF) were common in Irish language CMC. As discussed already, code switching was also used regularly for common interjections (e.g. oops, ooh, hmmm). The other two contexts for code switching were for (a) specialist vocabulary, usually associated with technology, and (b) some common everyday (often throwaway) words. These will now be explored further.

> Facebook message 27
> SeM: Ar fheabhas a Mhícheáil. Amach leat ar an hybrid anois le hullmhú dó!!!
> *Brilliant Michael. Head out on your hybrid now to prepare for it!!!*

> Blog comment 16
> DS: Níl an Ghaeilge ina *deciding factor* sna toghcháin, agus níorbh chosúil go mbeadh sí go brách.
> *Irish is not a deciding factor in elections, and it doesn't look like it will ever be.*

> Tweet 21
> Éinne i mBÁC ag iarraidh car pooling a dhéanamh go dtí an Scoil Earraigh don Phleanáil Teanga sa Ghaillimh Dé Luain? #dúbláil #comhoibriú
> *Anyone in Dublin interested in car pooling to go to the Spring School for Language Planning in Galway on Monday? #doubling #cooperation*

Uncommon, difficult or specialist terms often appeared in English across the three groups because the authors were either unfamiliar with the Irish equivalent or feared that the readers would be unfamiliar with it. The terms 'hybrid', 'deciding factor' and 'car pooling' have all been written in English, presumably because the Irish language equivalents did not come naturally to the authors. In the case of Tweet 21, the author may have used the English term for the benefit of the reader. In the same tweet, he created hashtags that used the Irish words for 'doubling' and 'cooperation', suggesting that he could also have devised an Irish language term for 'car pooling' if

he was so inclined. Switching to English words for specialist vocabulary is particularly common when discussing technology. Brand names or program names like Google Earth, Microsoft Works, MacBook, etc. appeared in English, as expected. However, English terms were not restricted to proper nouns. Much as English language acronyms relating to technology and web culture were common in Irish language CMC, phrases such as 'spam', 'web crawler', 'troll', 'data recovery', 'servers' and 'techies' were found in English across the three groups. Tweet 22 is a typical example of this phenomenon, where the non-Irish words "Beta" and "glitches" are dropped into otherwise Irish language content.

> Tweet 22
> Chuala mé gur seoladh Abair Leat inniú -- ní raibh sé réidh, i mo thuairimse. Fós ar leibhéal Beta, an-chuid glitches go fóill.
> *I heard that they launched [the website] Abair Leat today -- it wasn't ready, in my opinion. Still on Beta level, a lot of glitches still.*

In the ILT, there were many examples of the Irish language being adapted by core members to translate or at least to gaelicise some common technological terms from English – what might be called hybrid anglicisms.[93] This was often the case where vocabulary relating to Twitter itself was concerned. In Tweet 23 below the author has used the word "túit" and "ath-thúiteail" to describe a 'tweet' and the act of 're-tweeting'. Although her spelling of the word "túit" differs to the more standardised translation 'tuít',[94] both terms show an effort on behalf of the author to write monolingually in Irish without resorting to wholly English words. In Tweet 24 the verb 'googling' has not only been gaelicised but also adjusted to the genitive case according to the rules of Irish grammar, so it appears as "googlála". Although the double-o doesn't comply with Irish phonetics, in general the word has an Irish appearance that fits with the rest of the tweet. Similar examples from the ILT include "zoomáil" (zooming), "trendáil" (trending) and "blockáilte" (blocked), all of which break the rules of Irish spelling but were included in tweets. Though the words are obviously not of Irish origin, their use in this way acts to blend the non-Irish element

[93] The term 'hybrid anglicism' is used in Quebec to describe new words in Quebec French that are formed from a combination of an English word to which a French element is added. See the Quebec Office of the French Language's "Banque de dépannage linguistique": http://www.oqlf.gouv.qc.ca/francilettre/francilettreEX_20070914.html.
[94] From the online dictionary http://www.focal.ie.

into an otherwise Irish language structure. Such neologisms can be popularised and shared among users through their repeated use. In Tweet 25 one user is promoting the use of the morphological anglicism[95] "dí-lean" to a fellow Twitter user as a somewhat awkward Irish alternative for the English "unfollow". There were examples of standardised Irish translations of uncommon words referring to technology elsewhere in the ILT, including: "scáileán ard ghléine" (HD screen), "haischlib" (hashtag) and "leabharmharc" (bookmark). Although these are standardised Irish language terms they are not commonly used, and they demonstrate an investment on behalf of their authors in the maintenance of a high standard of Irish language content in their tweets.

Tweet 23
B'fhéidir nar coir dom an túit sin a ath-thúiteail, ach cur sé brón orm,& v mé ag iarraidh an brón a roinnt, go tógfaidh duine mo chroí arís
Maybe I shouldn't have re-tweeted that tweet, but it made me sad,& I wanted to share the sadness, that someone might take my heart again

Tweet 24
@username1 Breis googlála déanta agam . Feictear dom gur nath coitanta atá ann, eireaball eile leis go minic. http://URL1
@username1 I did more googling . I see that it is a common expression, another tail on it often. http://URL1

Tweet 25
@username1 @username2 Dí-lean. = Unfollow.
@username1 @username2 De-follow. = Unfollow.

It wasn't just for specialist or technological terms that code switching occurred. Common everyday words in English were peppered into Irish messages on blogs, Twitter and Facebook. This type of code switching differed from the practice of writing specialist terms in English, in that it would take little effort on behalf of the authors to write the words in Irish. Moreover, because they are common everyday words, it is unlikely that they are switching to English for the sake of their readers'

[95] The term 'morphological anglicism' is used in Quebec to describe new words in Quebec French that are literal translations of English forms where every element comes from the French language, but which reproduces, completely or partly, the image transmitted in original English.

comprehension. Rather, it demonstrates that Irish web users are comfortable communicating between the two languages in a casual manner. Switching to English in this instance was driven by style and choice, rather than compulsion. The practice of switching to English for everyday words was most common in the ILT, where the majority of users communicated in Irish less than half the time and where there were no rules governing language use. In general the ethos of the ILT was one of linguistic innovation and choice. However, code switching to common terms also occurred occasionally in comments posted on monolingually Irish blogs and messages on the *GA* wall. In Tweet 26 the author, a native Irish speaker, throws in the everyday phrases "great call" and "brilliant" into her otherwise Irish language tweet. She hasn't code-switched in this case because of a lack of an Irish language equivalent; rather, it is a stylistic feature of her discourse. In Blog comment 17, another native speaker uses the English word "sight" in an otherwise Irish language tweet. These English words have been incorporated into the syntax of Irish sentences, without interrupting the flow or meaning. Short English adverbs were particularly common in this context, with words like "so", "just", "really" and "sure" (see Facebook message 28) appearing in otherwise Irish sentences, perhaps mimicking the way in which these words are dropped naturally and casually into speech. In the context of speaking they 'roll off the tongue'; in the context of Irish language CMC they rolled off the keyboard.

Tweet 26
@username1 great call aréir bhreathnaigh mé ar Downton Abbey go dtí a 3.00am! Brilliant.
@username1 great call last night I watched Downton Abbey until 3.00am! Brilliant.

Blog comment 17
AOT: Seans go raibh an trá folamh ceart go leor… ní fhaca me sight di le tamall!
There's a chance the beach was empty alright… I haven't caught sight of her in a while!

Facebook message 28

BrF: Sure nach cuma, má tá sé ag déanamh jab maith mar mhéara ar London.

Sure it doesn't matter, if he's doing a good job as mayor of London.

There were no set ways of marking non-Irish words. Rather, non-Irish words and phrases were integrated into the text using different strategies at the behest of the individual authors. Outside of direct quotes, non-Irish words were commonly included in Irish language sentences without any graphic markers or spelling changes. For example, in the three messages above "great call", "Brilliant", "sight" and "Sure" appeared as unmarked standard English spellings in the same typography as the Irish words around them. In many cases, however, non-Irish words were altered, appearing differently from how they would normally appear in text. Sometimes non-Irish words were spelt differently to conform in part to the rules of Irish grammar. In other cases graphic markers (e.g. inverted commas, italics) were used to single out non-Irish loan words. These features were most common in the ILB, but there were examples from the other online groups also. The increased frequency in the ILB can be understood in the context of the general practice among bloggers of posting content that was as monolingually Irish as possible. When they resorted to using non-Irish words ILB members often adapted them in some way to Irish spelling, or marked them apart in various ways.

Blog comment 18

NI: Ní hé seo an chéad *pheace pole* a chonaic mé in áiteanna éagsúla.

This isn't the first peace pole *I've seen in different places.*

Tweet 27

@username1 Go hawesome. :)) Tá an pionta riachtanach!

@username1 Awesome. :)) The pint is essential!

Facebook message 29

LuM: bìonn sè ar an tsofa eadar mè fèin agus mo chailìn :-)

he's usually on the sofa between me and my girl :-)

In the above examples, the spelling of the non-Irish words "peace pole", "awesome" and "sofa" was tweaked to conform, in part, to Irish grammar. These English words have had letters added to them, to adapt them in some way to fit the grammatical rules of the Irish words around them. Without describing the grammar of Irish in detail, "peace pole" has been spelt with an extra 'h' as "pheace pole" because it is preceded by the Irish phrase "an chéad" (the first). Similarly "hawesome" has had a 'h' added because it is being treated as an Irish adjective beginning with a vowel should be treated following the Irish word "go". In standard Irish, feminine nouns beginning with 's' are preceded by the letter 't' when following the composite preposition "ar an" (on the). In Facebook message 29, the author has decided to apply this rule to the word "sofa", spelling it "tsofa", despite the word not being Irish and it not having any gender in the English language. These examples demonstrate some users' comfort in integrating non-Irish words into messages, as well as their determination to maintain an overall Irish language character to much of their content.

A more extreme version of integrating non-Irish words can be found in the following examples where the spelling is changed, sometimes beyond recognition, to conform to the phonetics of the Irish language. This was often done in a playful spirit, and was not a common feature of any group. In the following examples the terms "typos", "fair play" and "fuck 'em all" have been spelt using Irish phonetics as devised by the authors themselves, appearing as "taidhpos", "fér plé" and "focamál" respectively. The spelling of these quasi-Irish words may appear alien to non-Irish speakers, but when pronounced according to the phonetics of the Irish language they mimic their English spoken equivalents.

Blog comment 19
SFA: Cad a dhéanfaidh muid feasta gan taidhpos, …
What would we do without typos, …

Tweet 28
@username1 alt spéisiúil, fér plé :) cár fhan tú in éirinn?
@username1 interesting article, fair play :) where did you stay in ireland?

Facebook message 30
LuM: Agus scríobh mar a labhras tú agus focamál :-D
And write as you speak and fuck 'em all :-D

The above examples show the varying degrees to which Irish social media users attempted to blend non-Irish words into Irish language content. On the other hand, some users exhibited subtle strategies in singling out non-Irish words from the rest of their Irish language content. In the following examples, italics and/or inverted commas are used to mark the 'foreign' words and phrases. The objective is not to emphasise the words but, rather, to distinguish them from the rest of the – primarily Irish language – content. Italics were commonly used in the ILB. However, writing tweets in italics is impossible and writing wall posts on Facebook in italics requires using a separate application, so no examples of italicisation were found in the ILT or on *GA*.

Blog comment 20
ECB: Cuireann scéalta mar seo na *giggles* orm.
Stories like this give me the giggles.

Tweet 29
@username1 @username2 táim sínithe suas. Is maith an rud é go bhfuil laochra cosúil libh ag 'championing' an teanga! #gaeilge
@username1 @username2 i'm signed up. It is a good thing that heros like you are 'championing' the language! #irish

Blog comment 21
RAM: An chéad rud eile, chuala mé caint uathu ar ábhar an *'menopause'*.
*Next thing, I heard them talking about the '*menopause*'*.

Finally, Irish social media users exhibited a number of innovative strategies in their writing for translating Irish language material into other languages, and *vice versa*. This was most often undertaken to ensure that unfamiliar, modern or specialist words and phrases were not misinterpreted by the audience. However, there was also a sense that many users enjoyed being creative with translation, and discussions about

the use of new words and phrases were common. One approach was to write the translation immediately after the word/phrase, using brackets or a slash. The Irish terms "Faiche Stiabhna", "uibheagán" and "guthú" have been translated in the messages below as "Stephens Green" [sic], "omelette" and "V/O".

Tweet 30
@username1 tá siad i lár mBláth Cliath, in aice le Faiche Stiabhna (Stephens Green) nó an Dáil (Patrick Guilbaud)
@username1 they are in the centre of Dublin, beside Faiche Stiabhna (Stephens Green) or the Parliament (Patrick Guilbaud)

Blog comment 22
NI: Tá sé réasúnta bog, cosúil le huibheagán / *omelette* tirim.
It is reasonably soft, like a dry uibheagán / omelette.

Facebook message 31
AoG: Bhfuil a fhios ag duine ar bith caidé'n doigh a fhaigheann tú post guthú (V/O) ar Cartúin Gaeilge…? Tá mé ag cuardach post ann le fada an lá.
Does anyone know how you get a guthú (V/O) job on Irish Cartoons…? I've been looking for that sort of job for a long time.

Another approach was for the authors themselves to attempt to create a translation for the words or phrases. By including question marks, this approach appears to invite the audience to confirm the merits of the translation or to suggest more appropriate ones.

Blog comment 23
MMR: Nuair a bhí mé i mo pháiste óg, bhí líne i bpáirt (?party line) againn.
When I was a young child, we had a líne i bpáirt (?party line).

Blog comment 24

NI: […] níl sa teachtaireacht ach patrún ar "electronic reader board" (clár léitheoireachta leictreonach?).

[…] the message is only a pattern on an "electronic reader board" (clár léitheoireachta leictreonach?).

Tweet 31

@username1 Aye, like tá sé ceart go leor greim (a bunch?) bláthanna a cheannacht domh. ;)

@username1 Aye, like it is ok buying them greim (a bunch?) of flowers. ;)

The final approach to translation described here is without question unique to CMC and exhibits a highly innovative way of enabling readers to switch between languages. In the ILB, bloggers frequently created hypertext links from some words and phrases (viewed in the examples below underlined) to external websites. By clicking on the link, readers are brought to a dictionary definition, encyclopaedic entry or place name translation of the unfamiliar word(s). These differ from other hypertext links embedded within the ILB in that they are not created to draw attention to material on other websites or blog posts, but rather to help with the interpretation of the blog content itself. It enables the author to use unusual or specialist terminology without the risk of alienating their audience. The following three examples used links to an online Irish language dictionary (http://www.focal.ie), an online archive of Irish place names (http://www.logainm.ie), and an English language webpage from an online encyclopaedia (http://www.wikipedia.org) respectively.

Blog comment 25

ACA: Ansin, níl ann ach tualaing aige chun bheith i láthair an lae inniu.

Then, there's only the potential for it to be present today.

Blog comment 26

SFA: Is cuimhin liom freisin freastal ar an sochraid san tSnaidhm.

I also remember attending the funeral in Sneem.

Blog comment 27
RAM: An é comhairliú géiniteach atá i gceist agat?
Do you mean genetic counselling?

The advantage of translating the material in such an innovative way is that it makes no assumptions about the linguistic abilities of the reader. Rather, it gives the reader the option of referring to the translation where needed and, thus, avoids interrupting the flow of the content unnecessarily.

There was further evidence of the use of online Irish language resources in the choice of spelling in one blogger's translation of the English term 'tweet' (in the context of micro-blogging). SFA explained that he spelt his translation with reference to "abair.ie": a website that converts written words to speech (http://www.abair.ie).

Blog comment 28
SFA: Roghnaigh mé túít mar gur fhuaimnigh sé sin i gceart ag abair.ie
I chose [the word] túít because it sounded right on abair.ie

Each of the above examples shows a commitment on behalf of the authors to maintaining as monolingually Irish a space as possible in their interactions in the blogosphere, while not alienating less competent speakers.

6. DIALECT IN THE ILB

The Irish language is divided into three dialect groups – Munster Irish, Connacht Irish and Ulster (or Donegal) Irish – each with its own distinct vocabulary and pronunciation, as well as some grammatical variation. This study sought to explore how dialect impacted on online interaction in Irish, in the context of individuals communicating with each other through text on screen. We have already seen how individuals' dialects inspired syntactic and morphological variation. For example, Tweet 7 and Facebook message 8 above exhibited clues to the authors' dialects through their non-standard spelling, vocabulary and abbreviation. The dialect analysis that was conducted on the corpus of posts and comments on the ILB was based on lexicon identified as distinct to each dialect. The ILB was chosen because of the large size of the corpus of content extracted from the blogs. By classifying members of the

ILB according to these dialect identifiers, this study could show whether clusters formed around users who shared a common dialect, or indeed whether differences in dialects appeared to repel certain users from interacting with each other.

In traditional Irish language broadcast media, there are two different strategies for mixing or marking apart dialects. In servicing local Gaeltacht communities around the Republic of Ireland, the national radio station Raidió na Gaeltachta adopts a strategy of dividing their schedule according to the three dialect groups. The schedule alternates between programmes broadcast from studios in the Gaeltacht areas, with hosts and guests who most often speak the local dialect. Even news bulletins throughout the day feature separate news updates from Munster, Connacht and Ulster read by three newsreaders with corresponding dialects. In contrast, on the Irish language television station TG4, the predominant strategy is one of mixing dialects across programmes. In the evening their studio-based current affairs programmes include guests from around the country, while during the day their children's cartoons are likely to include a mix of characters with varying dialects. This study was interested in determining which of these strategies – one of separating and stratifying different dialects or one of mixing and mingling them – occurred in an online context.

71 individuals in the ILB (47%) were found to have dialect codes in their communications, 56 bloggers (82.4%) and 15 commenters (18.1%). In general commenters published far less content than bloggers and, thus, made up only a fraction of the ILB corpus. This is reflected in the low number of commenters identified as having dialect codes. Each blogger and commenter was given an individual dialect score. This was achieved by, firstly, scoring them one point for each of the dialect codes that appeared in their content and an extra point if that code featured three or more times in their content. Secondly, for each individual, the dialect with the highest score had the combined score of the other two dialects, if any, subtracted from it. This resulted in 54 individuals being assigned a single score corresponding to one dialect. The most significant scores for each dialect as well as the members' geographic locations are listed in Tables 34-36.

These scores would appear to support the choice of dialect codes for this study. The top score for each dialect (and the three highest scores in general) is held by individuals living in, or originally from, Gaeltacht areas in their identified dialect regions. Moreover, other high scores in the Connacht and Ulster dialects are held by individuals living in those provinces (ASF in Mayo, and CT in Armagh).

Surprisingly, two of the highest scorers (LIY and TSD) are self-described learners living outside the Gaeltacht (in Dublin and California respectively). This would suggest that their learning materials and/or tutors use a specific dialect.

Table 34. Connacht dialect scores.

Blog/Commenter	Dialect Score	Location
RAM	11	Donegal (outside Gaeltacht), but originally from Connemara Gaeltacht
LIY	7	Dublin
NI	7	Washington State, USA
CC	5	Unknown
ASD	4	Notre Dame, USA
ASF	3	Mayo

Table 35. Ulster dialect scores.

Blog/Commenter	Dialect Score	Location
AOT	10	Donegal Gaeltacht
TSD	8	California, USA
CT	4	Armagh
ACD	3	Unknown
TE	3	Galway city

Table 36. Munster dialect scores.

Blog/Commenter	Dialect Score	Location
IG	9	Cork Gaeltacht
CD2	4	Unknown
RNG	4	Germany

Figure 34 visualises the ILB according to the nodes' dialect scores, if any. Most nodes did not show tendencies towards the use of vocabulary associated with any particular dialect and, therefore, were not colour-coded. Where nodes have been coded, however, they indicate that there do not exist clusters of nodes who share a common dialect cut off from the rest of the network. Dialect does not appear to bond members, or exclude interaction between members, in any significant way. In fact the core of the network comprises a mixture of users across the three dialects interacting with each other. This would indicate that despite Irish language dialects having an impact on the lexicon social media users bring to the web, as well as the syntactic and

morphological variation arising from the distinct grammar and pronunciation associated with each dialect region, dialect does not block individuals from communicating with and understanding each other in a text-based environment online. These results would suggest that in participating in the ILB, members do not seek out a particular dialect, but rather use the ILB to participate in an Irish language network in general.

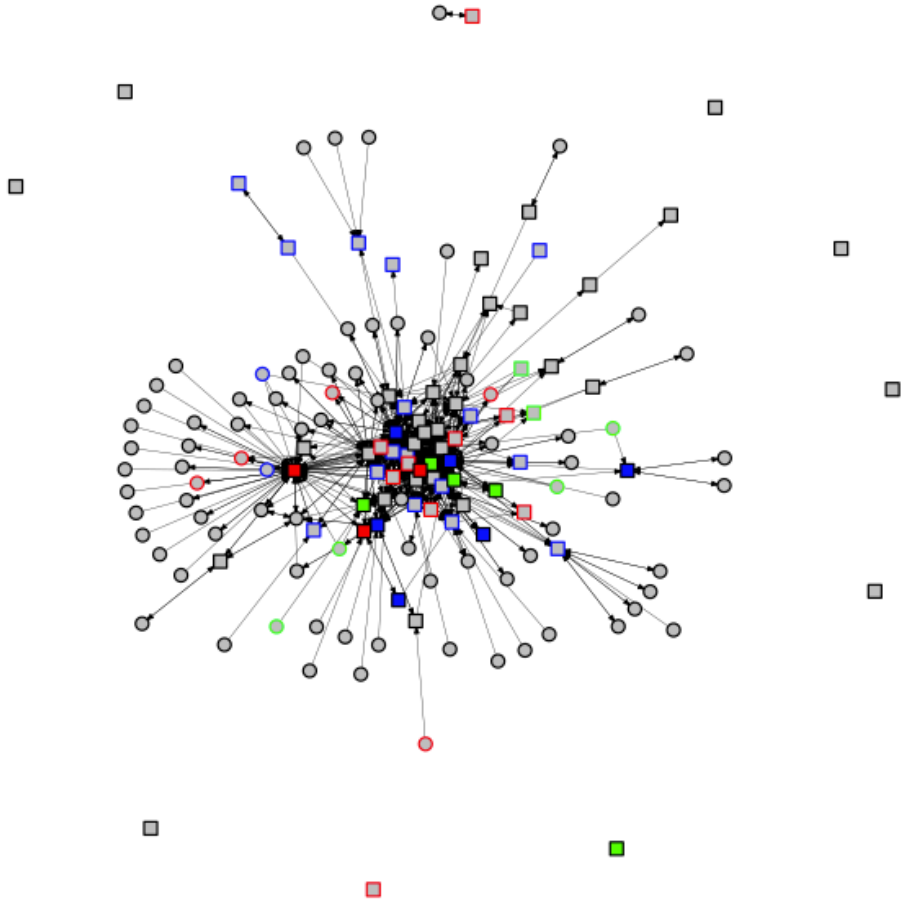


Fig. 34. Dialect in the ILB. Nodes are colour-coded as follows: fully red – Munster dialect score >2; red outline – Munster dialect score 1-2; fully blue – Connacht dialect score >2; blue outline – Connacht dialect score 1-2; fully green – Ulster dialect score >2; green outline – Ulster dialect score 1-2.

Having shown how the Irish language has been adapted in various ways by users of social media, and having identified different language features that marked a distinct language register in each group, this study now analyses whether the three groups may be called 'online communities'. Each group is now discussed individually.

D. ONLINE IRISH LANGUAGE COMMUNITIES

This study has established that there are distinct networks of Irish language users interacting with each other through social media online. Each network – typically composed of 100-300 active individuals – is comprised of a small cluster of densely connected core users where the majority of content is produced and interaction takes place, surrounded by a less well-connected and less prolific ring of peripheral users. We have seen how core users have adapted the Irish language to a text-based on-screen environment, adopting some features common to CMC in English (as well as other languages) while simultaneously innovating new ways of representing Irish speech on screen. The question remains whether these networks comprise 'online communities' according to the features used by sociolinguists to define communality on the web. The study will now look at each network and assess how they behave like communities according to the criteria laid out in Herring (2004).

1. THE IRISH LANGUAGE BLOGOSPHERE

In exploring the ILB using social network analysis and discourse analysis techniques, four of Herring's (2004) six criteria for online communities appear to be satisfied. Firstly, the statistics and social network analysis would seem to point to *active participation around a core group of regular participants*, although admittedly some members were more active than others. The core network members identified through their degree centrality were also the most prolific participants. Moreover, most core members were well established in the network with many having over three years experience of blogging in Irish. All of this points to a small cluster of individuals in regular contact with one another over time.

The distinct register peculiar to the ILB can be viewed as evidence of *shared norms and culture* among network members. This is seen, for example, in the innovative linguistic features for mimicking spoken language, abbreviating words and phrases, expressing emotion, and code switching, devised and shared between bloggers, and described above using Cherny's language features. Figure 35 identifies those bloggers and commenters that used two of the abbreviations common to the ILB – GRMA and ASG – during the data capture period. These abbreviations are concentrated in the ILB core, suggesting that those well-connected participants who

are in regular contact with one another are more likely to adopt features of the register.

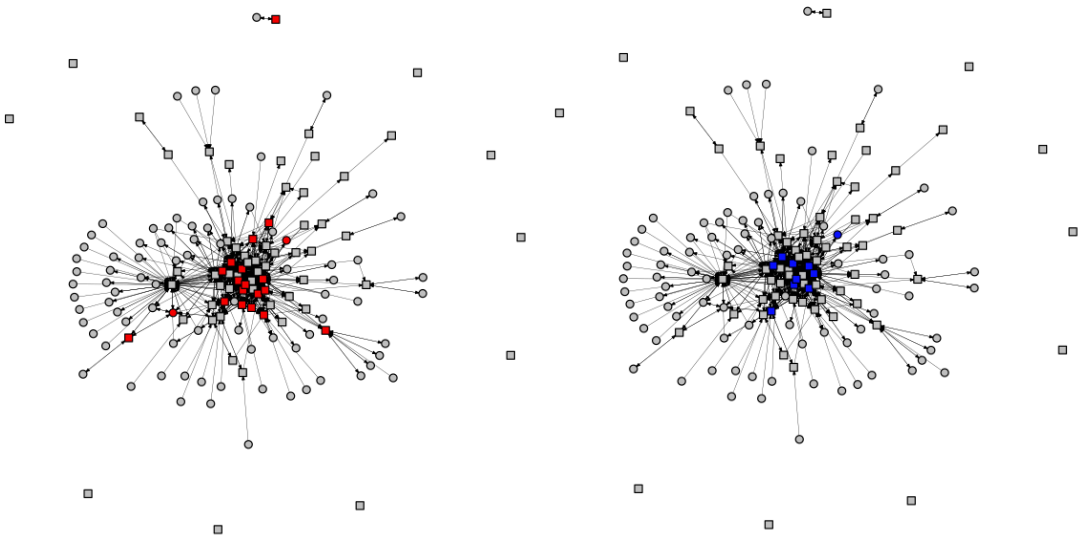


Fig. 35. Instances of the abbreviations GRMA (in red) and ASG (in blue) in the ILB.

Many of the features of the ILB register are best understood in terms of members' desire to maintain as monolingual an Irish space as possible. The practice of using hypertext links to offer translations of difficult or specialist words and phrases is one example peculiar to CMC and popular among core members of the ILB. Use of Irish language phonetic spellings of common interjections; using and creating Irish language abbreviations; offering personal translations into Irish of non-Irish words and phrases; and members correcting their own and others' spelling and grammatical errors occurred much more frequently in the ILB than in the other two groups. This is further evidence of the value core ILB members placed on the uniquely Irish language space created on and between their blogs. This points to another of Herring's criteria, that of *self-awareness* among group members that their group is unique. Further evidence of awareness among bloggers of a perceived wider network can be found in the way they address others in the ILB in the greetings and sign-offs in their posts. It was common in the ILB core to use the Irish phrase "a chairde" (equivalent to 'dear friends') to greet or address one's readership. Moreover, although the majority of bloggers write on personal themes, commonly in journal style blog entries, use of the second person plural to address thoughts or questions to a wider audience is common.

Posts addressed in these ways may not attract any comment in return. This suggests that bloggers communicate with an imagined community of readers in mind, however invisible and immeasurable that readership might be, and regardless of whether their messages or questions are responded to.

There is a generally supportive atmosphere between bloggers and commenters in the ILB, one in which reciprocal relationships are formed around the sharing of stories, poems, pictures, jokes, news and advice. Herring's criteria of *support and solidarity* is most clearly illustrated in the language routines described earlier through which ILB members expressed welcome, praise, encouragement, humour and sympathy. Short and formulaic responses enabled participants to register support with others with minimal time or effort. It could be argued that this type of support is rather shallow in its expression. However, it does suggest a desire among ILB members to offer tokens of support to others, despite the lack of social cues or face-to-face encounters, all in a potentially anonymous environment.

One particularly explicit expression of *solidarity* among ILB members, and one that also points to self-awareness among some ILB members of the distinctiveness of their community, occurred outside the ILB on the Irish language magazine website Beo! In an article in September 2012 entitled *'Blagadóireacht na Gaeilge, R.I.P.'* (Irish Blogging, Rest in Peace),[96] TG4 news editor Breandán Delap argued that the Irish language blogosphere had gone into terminal decline and needed to be 'resurrected'. To support his claim he gave examples of some of his favourite blogs that had in recent years fallen silent; he blamed flamers for discouraging people from blogging; he claimed that the lack of funding for blogging meant that journalists couldn't earn a living from it; and he claimed that Twitter was attracting some participants away from the blogosphere. The article received over 60 comments (by far the biggest response to any article on the website that year), including from some of the bloggers identified in this study. The response was, on the whole, critical of the article. Some accused the author of equating blogging to online journalism, and made it clear that their objective for blogging was a personal, and not journalistic, one – a pastime and not a profession. The ensuing lively debate spilled over into the blogosphere, where some bloggers were provoked into writing posts defending their blog practices.

[96] Published on the Beo! website: http://www.beo.ie/alt-blagadoireacht-na-gaeilge-rip.aspx.

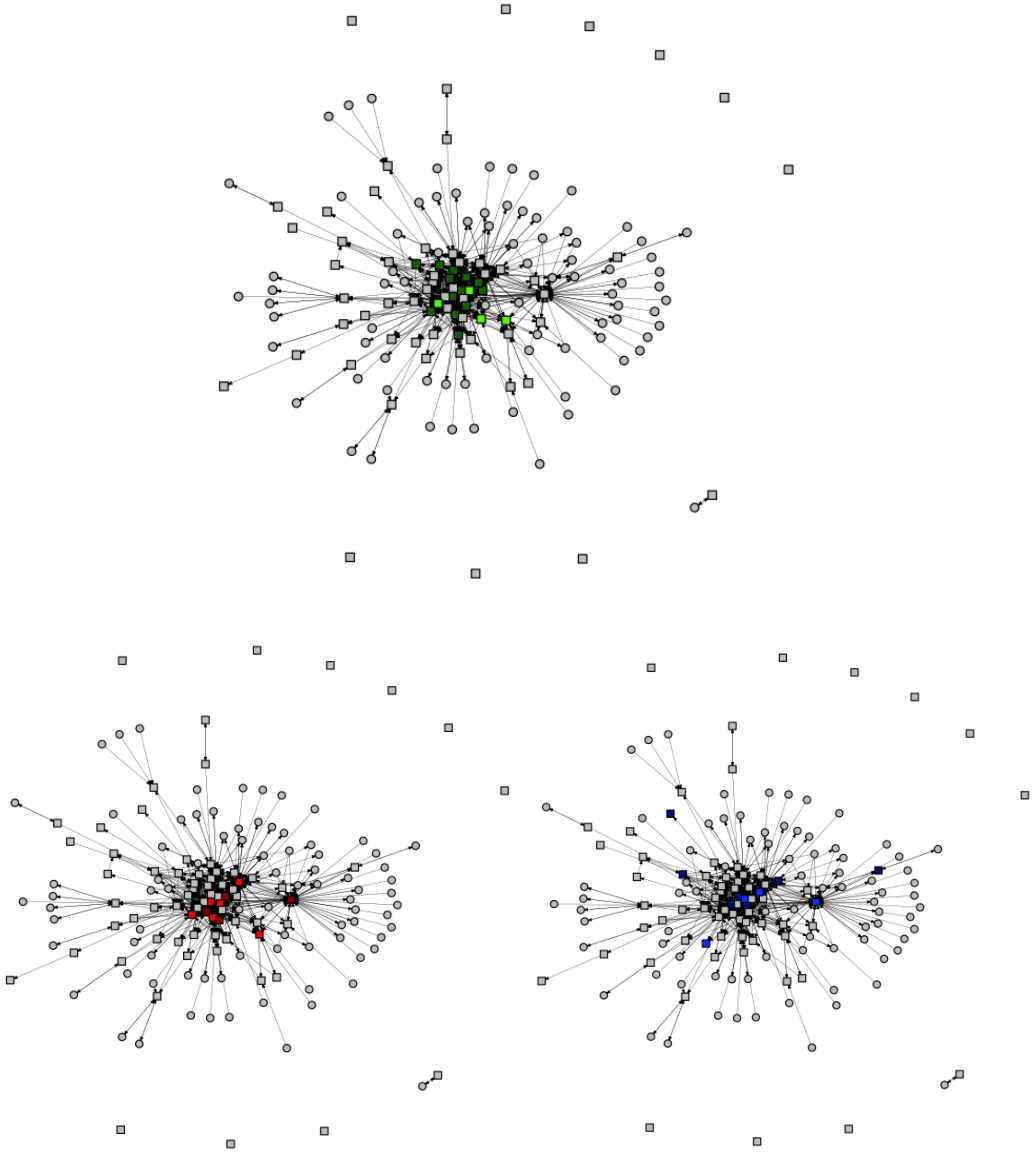


Fig. 36. Position of topics in the ILB: Personal blogs (in shades of green), Political blogs (in blue) and Current Events & News blogs (in red). Dark colours: primary theme/topic, light colours: secondary theme/topic.

The social network visualisations in figure 36 identify those blogs that are primarily personal journal style blogs, political blogs, or blogs focussing on news and current affairs respectively. It confirms that the core of the ILB is comprised primarily of personal blogs, displayed in shades of green. This would appear to support those criticising Breandán Delap's article for overly focusing on journalistic blogs in gauging the health of Irish language blogging. Although politics and news/current affairs are important topics to many core members of the ILB, they are fewer in number and less well-connected in the network. For most of the best-connected (and

most prolific) Irish language bloggers, blogging is a personal pursuit more focused on sharing thoughts and stories on a personal level than commenting on bigger themes and issues. When 'community blogs' are plotted, they are all located on the periphery of the network. Although they fulfil a role in informing their audiences about group news and events, they attract very little interaction from their readership.

This study has not addressed two of Herring's (2004) dimensions of online community in relation to the ILB: *conflict resolution* and *the emergence of roles and hierarchies*. The former was not included because of the scarcity of any moments of conflict between ILB members during the data capture period. The majority of blogs in the ILB were primarily personal or community blogs, posting messages that were unlikely to cause controversy. Furthermore, the most frequent commenters were other bloggers, who reciprocated back-and-forth on each other's blogs. The potential audience for blogs in a minority language like Irish is relatively small, and therefore bloggers might not want to risk alienating themselves by criticising or upsetting fellow bloggers. One of the political blogs (IG) at the core of the blogosphere did occasionally attract flaming by anonymous commenters. However, this was infrequent and was responded to with equal gusto by the blogger himself (an established journalist and newspaper editor). It was rare for other ILB members to become involved in diffusing any conflict on this blog.

The researcher did not identify any specific roles or hierarchies in the ILB. Social network analysis did identify one 'authority' in the network according to Herring *et al's* (2005) criteria. This blogger (IG, as described above) attracted relatively higher numbers of commenters, while commenting infrequently on other blogs. His unique position in the ILB was as a result of the frequent posts on politics, news and current affairs that attracted comment on his site from an audience that didn't engage beyond his blog. It does not, however, suggest any particular role that impacted interactions elsewhere in the ILB.

2. THE IRISH LANGUAGE TWITTERSPHERE

It is difficult to gauge if and how the ILT functions as an online community. The ILT, as it is imagined here, is made up of over 250 individuals who tweet multilingually, reading tweets from and writing posts to thousands of other Twitter users scattered across the globe. Unlike the ILB where the majority of bloggers

maintained monolingual Irish language sites, for many ILT members Irish was just one language among others through which they communicated. For the approximately two thirds of ILT members who tweeted in Irish less than half the time, the significance of an Irish language Twitter community may be lost in the mix of wider personal networks of followers and followees in other languages. Moreover, many Twitter connections are based on real-life acquaintances, much more so than in blogging. For many ILT members, Twitter may be just another medium through which to interact with friends and family through Irish, rather than a distinct online community.

That said, social network analysis does point to *active participation around a core group of regular participants* in the ILT, satisfying one of Herring's (2004) criteria for online communities. Although, admittedly, some members were more active than others in posting Irish language tweets, there appears to be a well-connected cluster of users at the core of the network in regular reciprocal interaction through Irish and referring directly to each other in their @replies, mentions and retweets. In fact, of the over 2,500 Irish language tweets collected from core members of the ILT, over 1,850 (or 74%) were directly addressed to at least one other Twitter user. This points to a well above average level of interactivity at the core of the network, when compared to the approximately 30% of tweets using the @ sign to address other users in Herring's (2009) sample of almost 37,000 tweets in multiple languages. The absence of separate clusters of users detached from the main network component suggests that Twitter users are connected through the Irish language to a wider network outside of just friends and families.

The distinct register peculiar to the ILT can be viewed as evidence of *shared norms and culture* among network members. There is a culture of non-standard language use that is made explicit in the wide range of linguistic and orthographic features used at the behest of the individual users. Through a range of syntactic and morphological variations, creative use of acronyms, play with modality, innovative code switching practices, and expressive use of hashtags, the core members of the ILT have adapted the Irish language to the specific communicative functions and structures of Twitter. The cumulative effect of all of these linguistic features is a creative and expressive use of the language that feels metamorphic, a space within which the expected rules of grammar and orthography are in a constant state of flux. Admittedly many of the practices derive from English language CMC, e.g. the use of

emoticons, the choice of phrases to acronymise, integration of @usernames and hashtags into the body of messages, etc. Many of the Irish language features explored above might have simply been adapted by individuals from their English language communication on Twitter, rather than emerging from group-wide participation. It would require further linguistic analysis across the ILT to determine how widely the use of specifically Irish language features – such as acronyms and interjections – was spread across the network.

The widespread use of common folksonomic hashtags based on Irish language words and phrases across the ILT suggests *self-awareness* among members of a distinct audience for their tweets. The common #gaeilge hashtag (i.e. the name of the language in Irish) is frequently used among ILT members in discussing a range of topics in the language. Hashtags, such as #lnag (*Lá na Gaeilge*, where people are invited to tweet in Irish for one day) or #onas12 (for the annual Irish language festival *Oireachtas na Samhna*), engage an undefined but specifically Irish language audience in interacting around individual online or offline events. These hashtags suggest that ILT members communicate with an imagined community of readers in mind, however invisible and immeasurable that readership might be. The organisation of offline social events – or tweetups – for Irish language Twitter users also suggests a perception that the community is distinct from others and deserving of its own targeted community events.

The researcher did not identify any specific *roles or hierarchies* in the ILT. One feature of Twitter is the power of celebrity accounts to attract high numbers of followers. While an international celebrity like Stephen Fry can attract over five million Twitter followers, the highest number of followers for an individual in the ILT was around 11,500 for the Irish boxer Bernard Dunne. Figure 37 shows the positions of the eleven ILT accounts that had 2,000 or more followers. These included two television personalities, a well-known journalist, a Northern Irish politician, a political party, and a regional radio station. Despite their large numbers of followers, most of these accounts were located peripherally in the network, and appear to engage directly with few of their followers through Irish. Despite their large followership affording them a vaulted status among other ILT members, they remain outside the core of activity.

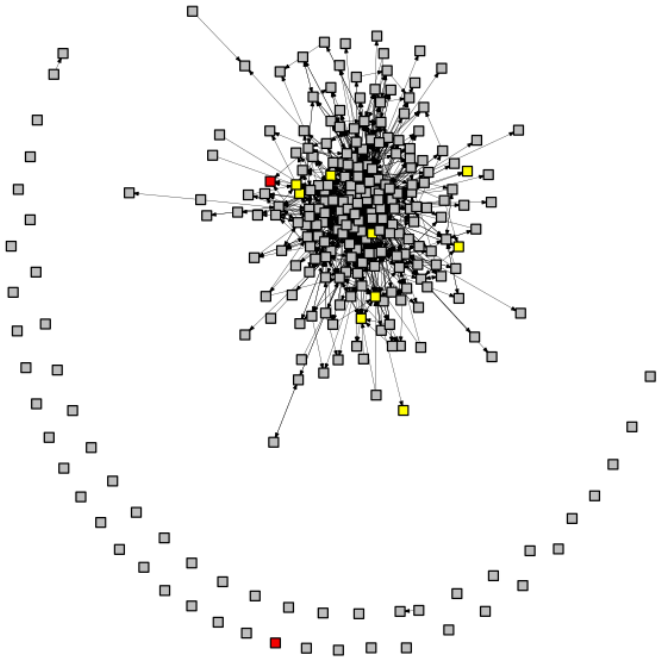


Fig. 37. ILT accounts with the highest numbers of followers.
Red = 10,000+ followers, yellow = 2,000+ followers.

3. THE *GAEILGE AMHÁIN* FACEBOOK GROUP

On the face of it, the *GA* Facebook group displayed many explicit features of an offline community. Firstly, many shared geographic proximity, with almost 80 members (or just over half of those whose location could be identified) situated in the province of Ulster in the north of Ireland. These were further concentrated in the neighbouring counties of Donegal and Derry. Secondly, some core members regularly met up with each other face-to-face and posted photographs of events – primarily in the Donegal Gaeltacht – where members were seen socialising. This was only one type of group member, however, and beyond these face-to-face interactions discourse analysis was needed to determine whether the interaction taking place on the group's wall characterised it as an online community according to Herring's criteria.

The statistics and social network analysis would seem to point to *active participation around a core group of regular participants*, although admittedly some members were more active than others. As in the other groups, core network members were also the most prolific participants. In contrast to the ILB and ILT, the *GA* group had a relatively short history, having been founded in 2011. Despite this, members

addressed each other informally, using each other's first names, and engaged in interaction around their interests, backgrounds and personal lives, signifying familiarity among regular contributors to the group.

In analysing the discourse of core members, one can identify a *hierarchy* within the group. In fact, of the three groups in this study, *GA* is the only one to have explicit positions of authority. These are occupied by the founder and group administrators, who may censor and delete messages that they deem inappropriate to the group's ethos or causing tension between members. These decisions are in turn endorsed or challenged by other group members.

There is a generally supportive atmosphere between *GA* members. 'Liking' messages and material posted by other members is commonplace. Reciprocal relationships are formed between members over time through the sharing of anecdotes, opinions, pictures and jokes. Anecdotes and pictures often refer to multiple members and their offline meetings. Herring's criteria of *support and solidarity* is most clearly illustrated among core members in the ways in which they rally around each other when the ethos of the group or the authority of the group's administrators is challenged by peripheral users. This most often occurs when the group's Irish only policy is challenged, and can be understood in terms of members defending their group's uniquely monolingual ethos.

Shared norms and culture in *GA* are most explicit in relation to the language rules of the group. From its title – *Gaeilge Amháin* (Irish only) – and the description in the 'About' tab, it is made clear to visitors and potential members that it is for communication through Irish only. Despite this, the group administrators and some of the more active members felt it necessary to post regular messages to the group wall reinforcing the Irish only ethos. These posts were commonly responded to and endorsed by other core members of the group. Another step in maintaining as monolingual an Irish space as possible was the deletion of non-Irish language content from the group wall by the administrators. On one occasion this caused conflict in the group, when one less active member began to question the practice of deleting links to English language material posted to the group wall. This member argued repeatedly that English language links should be allowed, that they inspired interaction between other members, and that the "foolish" practice of removing such posts turned the group into a "prison". Although the founder of the group (who had posted the message that prompted the attack) offered a confident counter-argument emphasising

the need for a strict Irish only policy, several core members of the group also got involved in the argument siding with the group's founder and endorsing the practice of deleting such links. Over time this built up a thread of conversation where member after member contributed their support for the practice.

The above instance is indicative of a group-wide strategy for *conflict resolution,* whereby the group's founder and administrators acted to resolve the conflict and restore order through their comments, which were in turn broadly endorsed by other core group members. There were a small number of examples of conflict during the study period, one in particular that had a lasting impact on the group. The conflict arose late one evening in a debate on the group wall where one member referred to native speakers being somehow superior to non-native speakers. The argument became heated, in particular between two members who were pitched against each other. The original post was deleted by the founder, who expressed regret that he was not online at the time (as he claimed that he would have put a stop to it). In the aftermath the original poster decided to take a break from contributing to the group, and the member who had been most insulted by his comments quit the group for good. This incident became the topic of later discussions on the group wall: some members claiming that the original post was insulting, others that the whole matter was being taken too seriously.

During the above incident one regular contributor described another Facebook group that was set up to encourage interaction through Scots Gaelic. This member argued that the absence of an explicit Scots Gaelic only policy contributed to most of the interaction taking place through English. The explicit support from most members of *GA*'s language policy, and their general adherence to it, demonstrates a certain *self-awareness* among group members that their group is unique. Unlike the ILB and ILT, whose networks were established in this study according to criteria set down by the researcher, *Gaeilge Amháin* is a bounded group with a specific number of clearly identifiable members. Individuals apply to become members of the group. Their membership can be granted and annulled. Admittedly the majority of these members are inactive. However, discourse analysis of active members showed members frequently referring to the wider group in their messages, as "GA", "ga", or "mo chairde GA" (my GA friends). One regular contributor went further, referring to this "suíomh álainn" (lovely site) as a "mionGhaeltacht" (mini-Gaeltacht).

Facebook message 32

AnK: is Gaeltacht bheag muid ag saothrú trí “Gaeilge Amháin”.

We are a small Gaeltacht operating through “Irish Only”.

Another regular contributor referred to *GA* as “an Ghaeltacht seo ar líne” (this online Gaeltacht). These users have adopted the term Gaeltacht as an established term for describing an Irish language community, albeit with very specific geographic and statutory definitions, and transferred it to an online context. Its use in this way demonstrates further self-awareness among some core members of the value of *GA* as an active Irish language community with a unique language policy.

VII DISCUSSION & CONCLUSIONS

A. KEY FINDINGS FROM THIS STUDY

What does the above social network and discourse analysis tell us about the vitality of the Irish language online? What does it reveal about where and how the language is used in CMC? And what findings does it present to those who might seek to promote further use of the language on the web? This study has shown that among the potentially hundreds of thousands of people with both web access and competency in the Irish language, only small networks of users clustered together publicly in significant numbers through a few popular social media sites to form what we might term online Irish language communities. Three of the most popular sites for interaction were the Irish language blogosphere (ILB), Irish language Twittersphere (ILT) and *Gaeilge Amháin* Facebook group (*GA*). The primary findings of this study are synopsised as follows:

- The types of discourse popular with Irish language users online were predominantly asynchronous, potentially anonymous, and (semi-)persistent.
- Of the seven primary and three secondary genres popular with Irish language users, blogging, micro-blogging and social networking sites were most popular. The most popular sites of interaction were Blogspot/Wordpress blogs, Twitter and Facebook.
- Each group in this study comprised a cluster of beween 150 and 300 active Irish language users.
- The groups shared a similar network structure: one main network component, with a core comprising the best-connected and most prolific users. Approximately one third of individuals in the ILB and *GA,* and one half of Irish language Twitter users, interacted with three or more other members through Irish. This core was surrounded by a periphery of less active users. Each group had some members who did not use the media interactively through Irish at all. Less than 10% of ILB and *GA* members were isolates, while just under 20% of ILT members were.
- The majority of users in each group were from outside the Gaeltacht, those rural communities designated by the Irish State as regions where the language

is said to survive as an everyday community language. In each group, where location could be identified, the largest cohort of users came from the island of Ireland outside the Gaeltacht. The ILB and ILT were particularly popular in Dublin, Belfast and Galway/Connemara; while *GA* members were concentrated in Ulster, particularly in counties Donegal and Derry. In addition, each group had members scattered across the globe.

- Women and young people were under-represented in all groups. Where gender could be identified, two-thirds or more of users in each group were male. Age data was not gathered for each group. However, it was clear from user profiles and from the discourse taking place that contributions from children and teenagers were rare.
- There were three distinct policies regarding code switching across the three groups. These ranged from the strictly imposed group-enforced 'Irish only' rule in *GA*, to a laissez-faire ethos in the Twittersphere where most users tweeted in more than one language. In the ILB, despite any explicit rules governing language use, most users chose to blog monolingually in Irish. Bloggers who posted bilingually were located on the periphery of the network. Where code switching to English did occur across the three groups it was most often for common interjections, or for uncommon or specialist terms, usually to describe technology.
- Members of each group were found to have adapted the Irish language to text-based CMC through features of syntactic and morphological variation; acronyms and abbreviation; play with modality; and language routines. Levels of non-standard use of the language varied between groups. This is best understood in terms of a spectrum of language variation. At one end, bloggers tended to compose their blog entries in fully formed monolingually Irish sentences, with standardised grammar, spelling and punctuation. Posts were commonly composed of paragraphs and comprised a beginning, middle and end. This contrasted markedly with the ILT, where abbreviated sentences with non-standard grammar, spelling and punctuation and frequent code switching were the norm. In *GA* too, grammar, spelling and punctuation were often non-standard although code switching was less common.

- The three groups displayed to varying degrees some or all of the criteria used by Herring (2004) to define online communities. All three groups, comprised active participation around a core group of regular participants. Discourse analysis among these core participants demonstrated shared norms and cultures unique to each group, as well as varying levels of self-awareness among members that their group was unique. Exchanges of support and solidarity were observed in the interactions between members of the ILB and *GA*. In addition, discourse in *GA* provided examples of strategies for conflict resolution, as well as the emergence of roles and hierarchies between group members, most noticeably in the enforcement of language rules by the founder and administrators.

B. INSTITUTIONAL SUPPORT FOR IRISH LANGUAGE COMMUNITIES ONLINE

The purpose of the above comparison between the online groups, as defined in this study, and the Gaeltacht, as it is defined by the State and popularly perceived by both residents and travel guide writers, is not to equate the two types of community. Nor is it intended to add any sense of legitimacy or authenticity to the Irish language interactions taking place through social media. However, it does challenge how we traditionally imagine communities of Irish language speakers, with their strong emphasis on place rather than contact. In doing so, it repositions the groups of individuals using social media to interact through Irish as unique language communities worthy of attention. Moreover, in the context of Irish being an endangered language, it poses questions about how these communities could and should be supported. In reporting the demise of the Irish language blogosphere, Breandán Delap (2011) suggested that sponsorship or financial support from state agencies (in particular *Foras na Gaeilge*, a state agency responsible for promoting the language) could help rescue Irish language citizen journalism on blogs. Some readers agreed that the lack of financial gain had contributed to the perceived decline in Irish language blogging. However, many bloggers commenting on the article argued that Foras na Gaeilge (or other agencies) should have no role in funding Irish language blogging. For them, blogging was a personal pursuit and not a focus for state investment, as these responses written by core members of the ILB express:

IG: Ar son Dé, ná lig don Fhoras aon láimh a chur in aice leis an mblagadóireacht. Más bocht suarach atáim agus mo leithéid, ar a laghad tá meas éigeann againn ar an rud a dheinimíd.
For God's sake, don't let Foras [na Gaeilge] near blogging. If I and my counterparts are poor and pitiful, at least we respect what we do.

CLR: Is ionann blagadóireacht agus breacadh agus roinnt smaointe le blagadóirí is léitheoirí eile. Níl éinne ag lorg "luach saothair", is é an scríbhneoireacht féin an "luach saothair" is mó le fáil uaidh.
Blogging involves expressing and sharing thoughts with other bloggers and readers. No one is looking for "reward", the act of writing itself is the biggest "reward" to be had from it.

The online communities in this study have coalesced on some of the world's most popular social media sites. The majority of members are individual users, uploading content and administering their accounts on a voluntary basis. Some blogs and Twitter accounts in this study were run by state-funded organisations, but these were few in number and peripheral in their networks. The communities in this study, therefore, survive and thrive independent of government support, without the grants and targeted investment enjoyed by Gaeltacht communities and Irish language organisations in the real world.

Those participating in social media through Irish do, however, benefit from some institutional support indirectly. This support is not financial; rather, through their web presence and the online services they provide a range of state and semi-state institutions contribute to the vitality of Irish language discourse online. In particular, third-level institutions and public service broadcasters played an important role in the vitality of the three online communities in this study.

Third-level institutions. Online language tools designed to help users produce content in different languages are made available by a range of providers – commercial, individual and institutional. Large commercial media companies can make a huge contribution to the vitality of minority languages online by releasing versions of their

web services in those languages. The availability of an Irish language version of Google Translate is a prime example. Where there are gaps in the services provided by profit-driven enterprises, universities commonly play a key role in developing free-to-acess language tools for public use. Language and/or computer science departments at third-level institutions are frequently involved in developing minority language resources online. Three Irish universities have combined their own internal linguistic and computer expertise with government funding to produce a range of online dictionaries and encyclopedias. University of Limerick has developed the all-Irish dictionary *An Foclóir Beag,*[97] which provides definitions and synonyms for its entries, as well as additional grammatical information about various forms associated with cases for nouns and conjugation for verbs. Dublin City University, through its Irish language department *Fiontar*, has developed a detailed online Irish-English/English-Irish dictionary at Focal.ie[98] that includes many specialist terms; a database of Irish placenames in Irish and English at Logainm.ie;[99] and a collection of Irish language biographies of individuals who were connected to the language at Ainm.ie.[100] Trinity College Dublin provides a service on its website Abair.ie[101] that allows users to type in Irish language words and hear an audio file of how they should be pronounced. All of the websites described above may be used by translators, authors, journalists and scholars who are producing Irish language content in traditional genres. For them, these websites replace the traditional function of a printed dictionary or encyclopedia. They are particularly significant, however, in an online context where users can use hypertext links in innovative new ways to refer directly to their content where needs be. The existence of such valuable resources in lesser-spoken languages like Irish shows the important role well-funded minority language departments in universities can play in providing a high quality network of language tools online for the wider language community of users and learners alike.

Public service broadcasting. Public service broadcasters play a vital role in supporting minority languages through their investment in television, radio and web content that might be unprofitable for commercial media organisations to produce.

[97] http://www.csis.ul.ie/focloir
[98] http://www.focal.ie
[99] http://www.logainm.ie
[100] http://www.ainm.ie
[101] http://www.abair.ie

Radio and television programmes can take on a life of their own on the web where their audiences can access additional content, upload their own material and interact on topics related to their favourite shows. This may take place on official websites created and moderated by the broadcaster or in the conversations that take place in social media elsewhere online beyond the broadcaster's control. A good example of both is the creation of a dedicated subsite on RTÉ's website for the television series *Bernard Dunne's Bród Club,* in which boxer and presenter Bernard Dunne encouraged viewers of the series to sign up to the website and commit to using what Irish they had. Those who signed up could take part in an online 'community' by posting messages to a forum, interacting with the programme's production team and viewers through Facebook and Twitter, entering competitions, uploading their own photo and video content, and discovering Irish language events and classes near them. By the end of the series the website boasted 1.2 million viewers on television, had reached almost 1.5 million people on Facebook and had almost 170,000 unique visitors to the programme's website.[102] Although the programme did not reach its desired goal of 100,000 individuals signing up to the campaign, it undoubtedly made a big impact both in traditional and new media terms.

Television and radio programmes frequently promote their own hashtags on Twitter to encourage their fans to comment on their content. This researcher analysed hashtag use among core members of the Welsh and Irish language Twitterspheres over one month in June 2012. The results showed that the names of television channels, programmes and personalities, in Welsh, Irish and English, were among the most popular. The hashtags #S4C, #ylle (for Welsh language TV programme *Y Lle*) and #bbcqt (for the BBC's English language TV programme *Question Time*) were among the most commonly tweeted in Welsh, while #TG4, #bbcgaeilge and various hashtags relating to both TG4's Irish language drama *Rásaí na Gaillimhe* and their factual series *Scéal na Gaeilge* were popular in Irish. This would indicate that traditional media content is regularly referenced and reinterpreted in the many conversations that take place online. The popular use of hashtags referring to TV and radio content demonstrates how traditional media can help form connections between social media users sharing opinions and experiences online. In the context of the relatively small networks of minority language users interacting online, it would

[102] Statistics from the official Bród Club website, sourced on 13/08/2012 from: http://www.rte.ie/brodclub/

appear that the health of traditional minority language media and the vitality of online communities of language users are inextricably linked.

Traditional broadcasters were also popular sources of retweets in the Irish and Welsh Twitterspheres. Retweets are used to endorse, promote or comment on content from elsewhere in the Twittersphere in a way that explicitly references the original content's author. Popular sources of retweets in Irish and Welsh included TV channels and programmes like @bbcblas and @sgorio, radio stations like @RnaGBAC and @RaidioRiRa, and TV and radio journalists and personalities like @MaireTNuacht, @silliebee and @vaughanroderick. Messages from these sources and others are repeated and referred to by other members of the Twittersphere building up conversations between Irish/Welsh Twitter users over time, further contributing to the vitality of the languages in that web genre. Thus, the efforts of traditional media broadcasters, programme makers and personalities in publishing interesting and engaging minority language content online helps promote use of the language among amateur consumer-producers of content elsewhere on the web.

C. IN CONCLUSION

Irish is one of the languages of the web. Like other living languages in the developed world it is being used interactively on a range of web genres and sites. The number of Irish speakers quoted in census returns is not reflected in the vitality of the language online, however. Despite the hundreds of thousands of people who claim competency in Irish and who have access to the web, regular interaction in public forums is confined to small groups on just a few sites. Notwithstanding their small number and size, these online networks are worthy of study, as they represent a new type of Irish language community. Users are pioneering new ways of using the Irish language in these spaces, and they are not doing so alone. Group practices and registers have emerged through routine interaction between regular users over time.

Perhaps the most significant finding from this study is how geographically spread the online communities are; specifically, how poorly represented the Gaeltacht is in its share of participants. They are communities without propinquity, where members are bonded by shared interest rather than geography. Much like the concept of 'localities' in Castells' writings, the traditional notion of an Irish language community

has become "disembodied from its cultural historical, geographic meaning and reintegrated into functional networks" (Castells, 1996: 365) in cyberspace. Within these networks the language is used innovatively and in socially meaningful ways, such that the well-connected and prolific participants at the core of these networks can be characterised as online communities.

In his poem *Fiabhras* – Fever, the Irish language poet Seán Ó Ríordáin (1964) describes his slow drift into the after-life: *Tá ceantar ag taisteal ón spéir / Tá comharsanacht suite ar mo mhéar* – A locality is forming in the ether / a neighbourhood perches on my finger.[103] He may well have been describing the communities of Irish language users that have emerged in cyberspace, through their mouses and keyboards, in new forms of online discourse and social media. While others may predict language death and decline, this online activity generated by hundreds of users from within and beyond the languages' traditional heartlands points to a new form of language community in evolution. As soon as this study is printed it becomes out of date. Since the data capture period, the Irish language online landscape has already changed dramatically with the launch of AbairLeat, a specifically Irish language social networking site. It remains to be seen the impact of that launch. Indeed, it remains to be seen how the online communities at the focus of this study, and the ways in which they use the Irish language, evolve over time.

[103] Translation by Greg Delanty, *Poetry Ireland Review*, Issue 84.

VIII A NOTE ON RESEARCH ETHICS

This study adhered to the guidelines and processes laid out in Cardiff University's Research Governance Framework.[104] Throughout this study the researcher also referred to the 'Ethics Guide'[105] of the Association of Internet Researchers. While accepting that there are grey areas in the ethics of online research, the researcher took a number of steps to protect the privacy and identities of the research subjects. Firstly, discourse was only sourced from publicly accessible forums, and all samples of discourse given in the study were available to view by any member of the public with internet access, without the need for log-ins or subscriptions. Secondly, the authors of the blog posts/comments, tweets and Facebook messages quoted in this study have had their names anonymised, without exception. Moreover, the names of individuals referred to in those messages have been anonymised. Thirdly, although the researcher is an active blogger and a friend of the *Gaeilge Amháin* Facebook group, he has never prompted any members of these groups or solicited information from them about their web habits.

[104] Available for download here: http://www.cardiff.ac.uk/racdv/resgov/CU%20Research%20Governance%20Framework%20Final%20senate%20v4.0%20230610.doc

[105] Available here: http://aoir.org/reports/ethics2.pdf, last accessed 10/03/2013.

BIBLIOGRAPHY

Anderson, B. (1983), *Imagined Communities*. London: Verso.

Anderson, J. F., Beard, F. K. & Walther, J. B. (2010), "Turn-taking and the local management of conversation in a highly simultaneous computer-mediated communication system" in Language@Internet, Volume 7. Retrieved from web on 10/10/2013: http://www.languageatinternet.org/articles/2010/2804

Androutsopoulos, J. (2010), "Localizing the Global on the Participatory Web" in Coupland, N. (Ed.), *The Handbook of Language and Globalization*. Chichester: Wiley-Blackwell.

Androutsopoulos, J. (2011), "Language change and digital media: A review of conceptions and evidence" in Tore, K. & Coupland, N. (Eds.) *Standard languages and language standards in a changing Europe*. Oslo: Novus.

Arendt, H. (1958), *The human condition.* Chicago: University of Chicago Press.

Bakardjieva, M. (2007), "Virtual Togetherness: An everyday-life perspective" in Bell, D. & Kennedy, B. (Eds.) *The Cybercultures Reader* (2nd Edition). Abingdon, Oxon: Routledge.

Baron, N. S. (2010), "Discourse structures in Instant Messaging: The case of utterance breaks" in Language@Internet, Volume 8. Retrieved from web on 10/10/2013: http://www.languageatinternet.org/articles/2010/2651

Bauman, R. (2001), "The ethnography of genre in a Mexican market: form, function, variation" in Eckert, P. & Rickford, J. R. (Eds.), *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Baym, N. (1998), "The Emergence of On-Line Community" in Jones, S. (Ed.), *Cybersociety 2.0: Revisiting Computer-Mediated Communication and Community*. Thousand Oaks, CA: Sage Publications.

Baym, N. (2010), *Personal Connections in the Digital Age*. Cambridge: Polity Press.

Bellah, R., Madsen, R., Sullivan, W., Swidler, A. & Tipton, S. (1985), *Habits of the heart: Individualism and commitment in American life*. Berkeley: University of California Press.

Bergs, A. (2006), "Analyzing online communication from a social network point of view: questions, problems, perspectives". Published online: http://www.internet.de. Retrieved 3 February, 2011.

Bilson, A. J. (1995), "Get into the groove: Designing for participation" in *Interactions*, Vol. 2, No. 2.

Blood, R. (2002), *The weblog handbook: Practical advice on creating and maintaining your blog*. Cambridge, MA: Perseus Publishing.

Boissevain, J. (1979), 'Network analysis: a reappraisal' in *Current Anthropology,* Vol. 20.

Bourke, A. (2006), "Re-imagining the Gaeltacht: Maps, Stories, and Places in the Mind" in Wyndham, A. H. (Ed.), *Re-imagining Ireland*. Charlottesville, VA: University of Virginia Press.

Britain, D. & Matsumoto, K. (2005), "Language, Communities, Networks and Practices" in Bell, M. J. (Ed.), Clinical Sociolinguistics. Malden MA: Blackwell Publishing.

Butts, C. T. (2008), "Social network analysis: A methodological introduction" in *Asian Journal of Social Psychology*, Vol. 11, No. 1.

Castells, M. (1996), *The Rise of the Network Society. Volume 1 of The Information Age: Economy, Society and Culture*. Oxford: Blackwell.

Castells, M. (2007), "Communication, power and counter-power in the network society" in *International Journal of Communication*, Vol. 1, No. 1.

Castells, M. (2010), *The Rise of the Network Society* (2nd Edition). Oxford, UK: Blackwell.

Cherny, L. (1999), *Conversation and Community, Chat in a Virtual World*. Stanford, CA: CSLI Publications.

Climent, S., Moré, J., Oliver, A., Salvatierra, M., Sànchez, I., Taulé, M. & Vallnamya, L. (2003), "Bilingual newsgroups in Catalonia: A challenge for machine translation" in *Journal of Computer-Mediated Communication*, Vol. 9, No. 1.

Cohen, A. (1985), *The symbolic construction of community*. New York: Routledge.

Coupland, N. (1988), *Dialect in Use: sociolinguistic variation in Cardiff English*. Cardiff: University of Wales Press.

Cronin, M. (2005), *An Ghaeilge san Aois Nua/Irish in the New Century*. Dublin: Cois Life Teoranta.

Crystal, D. (1999), *A Dictionary of Language, Second Edition*. Chicago: The University of Chicago Press.

Crystal, D. (2000), *Language Death*. London: Routledge.

Crystal, D. (2006), *Language and the Internet*, *second edition*. Cambridge: Cambridge University Press.

Crystal, D. (2009), *The Future of Language*. London: Routledge.

Crystal, D. (2011), *Internet Linguistics: A Student Guide*. Abingdon, Oxon: Routledge.

Csárdi, G. & Nepusz, T. (2006), "The igraph software package for complex network research" in *InterJournal Complex Systems,* 1695.

CSO [Central Statistics Office] (2011), "Information Society and Telecommunications in Households 2009-2011". Dublin: Stationery Office. Published online at:
http://www.cso.ie/en/media/csoie/releasespublications/documents/informationtech/2011/
isth2009-2011.pdf.

Cunliffe, D. (2007), "Minority Languages and the Internet: New Threats, New Opportunities" in Cormack, M. & Hourigan, N. (Eds.), *Minority Language Media: Concepts, Critiques and Case Studies*. Clevedon, UK: Multilingual Matters.

Cunliffe, D. & Honeycutt, C. (2008), "The Blogiadur – a community of Welsh-language bloggers" in Sudweeks, F., Hrachovec, H. & Ess, C. (Eds.), *Proceedings of the 2008 Cultural Attitudes towards Technology and Communication (CaTaC) Conference.* Murdoch, Australia: Murdoch University Press.

Cunliffe, D. & Honeycutt, C. (2010), "The use of the Welsh language on Facebook: an initial investigation" in *Information, Communication and Society,* Vol. 13, No. 2.

Daming, X., Xiaomei, W. & Wei, L. (2008), "Social Network Analysis" in Wei, L. & Moyer, M. G. (Eds.), *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Oxford: Blackwell Publishing.

Danet, B. & Herring, S. (2007), *The Multilingual Internet: Language, Culture, and Communication Online*. Oxford: Oxford University Press.

Dawson, C. (2010), "Change of Ownership" in Clark, D. & Jarazo-Álvarez, R. (Eds.), *In the Wake of the Tiger: Irish Studies in the Twenty-First Century*. Olieros, Spain: Netbiblo.

December, J. (1993), "Characteristics of oral culture in discourse on the Net", unpublished paper cited and summarised by Benson, R. (1994), "Telecommunications and society: a review on the research literature on computer-mediated communication", Berkeley, CA: University of California, Berkeley Roundtable on the International Economy, Compuscript. Available online at: http://www.december.com/john/papers/pscrc93.txt

Deere, A. (2011), "Towards the Measurement of Minority Language Presence on the World Wide Web", PhD submitted to the University of Glamorgan. Wales: University of Glamorgan.

Delap, B. (2011), "Blagadóireacht na Gaeilge, R.I.P." [Irish Blogging, Rest in Peace], published on the Beo! website: http://www.beo.ie/alt-blagadoireacht-na-gaeilge-rip.aspx.

Donath, J. (1999), "Identity and deception in the virtual community" in Smith, M. A. & Kollock, P. (Eds.), *Communities in cyberspace*. London: Routledge.

Durham, M. (2003), "Language choice on a Swiss mailing list" in *Journal of Computer-Mediated Communication*, Vol. 9, No. 1.

Elias, N. (1974), "Foreword" in Bell, C. & Newby, H. (Eds.), *The sociology of community: A selection of readings*. London: Cass.

Fleming, A. & Debski, R. (2007), "The Use of Irish in Networked Communications: A Study of Schoolchildren in Different Language Settings" in *Journal of Multilingual and Multicultural Development*, Vo. 28, No.2.

Freeman, L. C. (1979), 'Centrality in social networks: Conceptual clarificiation' in *Social Networks*, Vol. 1.

Fruchterman, T. M. J. & Reingold, E. M. (1991), "Graph Drawing by Force-directed Placement" in *Software – Practice and Experience,* Vol. 21, No. 11.

Garton, L., Haythornthwaite, C. & Wellman, B. (1997), "Studying Online Social Networks" in *Journal of Computer-mediated Communication*, Vol. 3, No. 1.

Georgakopoulou, A. (2011) "'On for drinkies?': Email cues of participant alignments" in Language@Internet, Volume 8. Retrieved from web on 10/10/2013: http://www.languageatinternet.org/articles/2011/georgakopoulou

Gerrand, P. (2007), "Estimating Linguistic Diversity on the Internet: A Taxonomy of Pitfalls and Paradoxes" in *Journal of Computer-Mediated Communication*, Vol. 12.

Gibson, D., Kleinberg, J. & Raghavan, P. (1998), "Inferring Web communities from link topology" from *Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia*.

Giles, H., Bourhis, R. Y. & Taylor, D. M. (1977), "Towards a theory of language in ethnic group relations" in Giles, H. (Ed.), *Language, ethnicity, and intergroup relations*. New York: Academic Press.

Granovetter, M. S. (1973), 'The Strength of Weak Ties' in *American Journal of Sociology,* Vol. 78, No. 6.

Greffenstette, G. & Nioche, J. (2000), "Estimation of English and non-English Language Use on the WWW" from *Proceedings of RIAO'2000 (Content-Based Multimedia Information Access)*, Paris, April 12-14, 2000.

Guinovart, X. G. (2003), "A lingua galega en Internet [The Galician language on the Internet]" in Bringas, A. & Martín, B. (Eds.), *Nacionalismo e Globalización: Lingua, Cultura e Identidade*. Vigo, Spain: Universidade de Vigo.

Gunn, M. (1997), "A Gaeltacht on the Internet" in *UCD News.* Dublin: University College Dublin.

Haythornthwaite, C. (2002), "Strong, Weak, and Latent Ties and the Impact of New Media" in *The Information Society*, No. 18.

Herring, S. C. (Ed.) (1996), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives.* Amsterdam: Benjamins.

Herring, S. C. (2001), "Computer-mediated Discourse" in Schiffrin, D., Tannen, D. & Hamilton, H. (Eds.), *The Handbook of Discourse Analysis*. Oxford: Blackwell Publishers.

Herring, S. C. (2004), "Computer-mediated discourse analysis: An approach to researching online behavior" in Barab, S., Kling, R. & Gray, J. (Eds.) *Designing for Virtual Communities in the Service of Learning*. Cambridge, UK: Cambridge University Press.

Herring, S. C. (2007), "A Faceted Classification Scheme for Computer-Mediated Discourse" in *Language@Internet*, Vol. 1. Published online: http://www.languageatinternet.de. Retrieved February 2011.

Herring, S. C. (2009), "Beyond Microblogging: Conversation and Collaboration via Twitter", in *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences (HICSS-42)*. Los Alamitos, CA: IEEE Press.

Herring, S. C. (2013), "Discourse in Web 2.0: Familiar, reconfigured, and emergent" in Tannen, D. & Tester, A. M. (Eds.) *Georgetown Univeristy Round Table on Languages and Linguistics 2011: Discourse 2.0: Language and new media.* Washington, DC: Georgetown University Press.

Herring, S. C., Kouper, I., Paolillo, J., Scheidt, L. A., Tyworth, M., Welsch, P., Wright, E. & Yu, N. (2005), "Conversations in the Blogosphere: An Analysis 'From the Bottom Up'" from *Proceedings of the 38th Hawai'i International Conference on System Sciences*.

Herring, S. C., Paolillo, J., Ramos-Vielba, I., Kouper, I., Wright, E., Stoerger, S., Scheidt, L. A. & Clark, B. (2007), "Language Networks on LiveJournal" from

Proceedings of the Fortieth Hawai'i Iternational Conference on System Sciences (HICSS-40). Los Alamitos: IEEE Press.

Hickey, R. (2011), *The Dialects of Irish, Study of a Changing Landscape*. Berlin: De Gruyter.

Hine, C. (2000), *Virtual Ethnography*. London: SAGE.

Huberman, A., Romero, D. M. & Wu, F. (2009), "Social networks that matter: Twitter under the microscope" in *First Monday*, Vol. 14, No. 1, 5 January, 2009. Sourced on 27/03/2012 from: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2317/2063.

Hummon, D. (1990), *Commonplaces: Community ideology and identity in American culture.* New York: State University of New York Press.

Jarbou, S. & al-Share, B. (2012) "The effect of dialect and gender on the representation of consonants in Jordanian chat" in Language@Internet, Volume 9. Retrieved from web on 10/10/2013: http://www.languageatinternet.org/articles/2012/jarbou

Joinson, A. N. (2003), *Understanding the Psychology of Internet Behaviour: Virtual Worlds, Real Lives*. Basingstoke: Palgrave MacMillan.

Johnson, N. C. (1997), "Making space: Gaeltacht policy and the politics of identity" in Graham, B. (Ed.), *In Search of Ireland: a cultural geography*. Routledge: London.

Jonassen, C. (1959), "Community typology" in Sussman, M. (Ed.), *Community structure and analysis*. Thousand Oaks, CA: Sage.

Judge, J., Ní Chasaide, A., Ní Dhubhda, R., Scannell, K. P. & Uí Dhonnchadha, E. (2011), *The Irish Language in the Digital Age/An Ghaeilge sa Ré Dhigiteach*. Available online at: http://www.meta-net.eu/whitepapers/e-book/irish.pdf.

Kaplan, A. M. & Haenlein, M. (2010), "Users of the world, unite! The challenges and opportunities of social media" in *Business Horizons,* Vol. 53, No. 1.

Kelly-Holmes, H. (2006), "Irish on the World Wide Web: Searches and sites" in *Journal of Language and Politics*, Vol. 5, No. 2.

Kelly-Holmes, H. (2011), "Sex, lies and thematising Irish. New media, old discourses?" in Kelly-Holmes, H. & Milani, T. M. (Eds.), *Thematising Multilingualism in the Media*. Amsterdam: John Benjamins Publishing Co.

Kirshenblatt-Gimblett, B (1996), "The Electronic Vernacular" in Marcus, G. (Ed.), *Connected: Engagements with Media*. Chicago: University of Chicago Press.

Kytölä, S. and Androutsopoulos, J. (2012), "Ethnographic Perspectives on Multilingual Computer-Mediated Discourse, Insights from Finnish Football Forums on the Web" in Gardner, S. & Martin-Jones, M. (Eds.) *Multilingualism, Discourse, and Ethnography*. London: Routledge.

Langer, S. (2001), "Natural languages and the world wide web" in *Bulag. Revue Annuelle*. France: Presses Universitaires Franc-Comtoises. Online: http://www.cis.uni-muenchen.de/people/langer/veroeffentlichungen/bulag.pdf

Lave, J. & Wenger, E. (1991), *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Lazer, D., Pentland, A., Adamic, L., Sinan, A., Barabási, A-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., Van Alstyne, M. (2009), "Computational Social Science" in *Science*, Vol. 323.

Lenihan, A. (2011), "'Join Our Community of Translators': Language Ideologies and/in Facebook" in Thurlow, C. & Mrocsek, K. (Eds.), *Digital Discourse: Language in the New Media*. Oxford: Oxford University Press.

Lockard, J. (1997), "Progressive politics, electronic individualism and the myth of virtual community" in Porter, D. (Ed.), *Internet culture*. New York: Routledge.

Lysloff, R. (2003), "Musical Community on the Internet: An On-line Ethnography" in *Cultural Anthropology*, Vol. 18, No. 2. Washington, D.C.: American Anthropological Association.

Mac Giolla Chríost, D. (2005), *The Irish Language in Ireland: From Goídel to Globalisation*. Abingdon, Oxon: Routledge.

Mac Giolla Chríost, D. (2012), *Jailtacht: The Irish Language, Symbolic Power and Political Violence in Northern Ireland, 1972-2008*. Cardiff: University of Wales Press.

Marlow, C. (2004), "Audience, structure and authority in the weblog community" from the International Communication Association Conference, May, 2004, New Orleans. Retrieved from http://rockngo.org/wp-content/uploads/mt/archives/ICA2004.pdf

Marsen, S. (2006), *Communication Studies*. Basingstoke, UK: Palgrave MacMillan.

Mas I Hernàndez, J. (2003), "La salut del català a Internet [The state of health of Catalan on the Internet]. Online: http://www.softcatala.org/articles/article26.htm Retrieved October 11, 2006.

McKenna, K. & Bargh, J. (1998), "Coming Out in the Age of the Internet: Identity "Demarginalization" Through Virtual Group Participation" in *Journal of Personality and Social Psychology*, Vol. 75, No. 3. Ithaca, NY: SPSP.

Mesthrie, R. (1999), *Introducing sociolinguistics*. Edinburgh: Edinburgh University Press.

Meyrowitz, J. (1986), *No Sense of Place, the Impact of Electronic Media on Social Behavior*. Oxford: Oxford University Press.

Miller, D. & Slater, D. (2000), *The Internet: an ethnographic approach*. London: Berg.

Milroy, J. (1992), *Linguistic Variation and Change*. Oxford: Blackwell Publishing.

Milroy, L. & Gordon, M. (2003), *Sociolinguistics: Methods and Interpretation*. Oxford: Blackwell Publishing.

Milroy, J. & Milroy, L. (1992), "Network Structure and Linguistic Change" in Milroy, J., Linguistic Variation and Change: On the Historical Sociolinguistics of English. Oxford: Blackwell Publishing.

Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P. & Bhattacharjee, B. (2007), "Measurement and Analysis of Online Social Networks" in *IMC '07*, October 24-26, 2007, San Diego, California, USA.

Mitra, A. (2000), "Virtual Commonality: Looking for India on the Internet" in Bell, D. & Kennedy, B. (Eds.) *The Cybercultures Reader*. Abingdon, Oxon: Routledge.

Moseley, C. (Ed.) (2010), *Atlas of the World's Languages in Danger, 3rd edition*. Paris: UNESCO Publishing. Online version: http://www.unesco/org/culture/en/endangeredlanguages/atlas

Ní Chartúir, D. (2002), *The Irish Language: An Overview and Guide*. New York: Avena Press.

Nic Pháidín, C. (2008), *A New View of the Irish Language*. Dublin: Cois Life Teoranta.

Ó Broin, B. (2010), "Schism fears for Gaeilgeoirí" in *Irish Times*, 16/01/2010. Dublin: Irish Times.

Ó Cuív, B. (1980), *Irish Dialects and Irish-Speaking Districts, Three Lectures by Brian Ó Cuív*. Dublin: Dublin Institute for Advanced Studies.

Ó Giollagáin, C., Mac Donnacha, S., Ní Chualaín, F., Ní Shéaghdha, A., O'Brien, M. (2007), "Comprehensive Linguistic Study of the Use of Irish in the Gaeltacht". Dublin: The Department of Community, Rural and Gaeltacht Affairs.

Ó Laoire, M. (2005), "The Language Planning Situation in Ireland" in *Current Issues in Language Planning*, Vol. 6, No. 3. Clevedon, UK: Multilingual Matters Ltd.

O'Rahilly, T. F. (1972), *Irish Dialects Past and Present.* Dublin: Dublin Institute for Advanced Studies.

Ó Riagáin, P. (1992), *Language Maintenance and Language Shift as Strategies of Social Reproduction. Irish in the Corca Dhuibhne Gaeltacht 1926-86*. Dublin: Institiúd Teangeolaíochta Éireann.

Ó Riagáin, P. & Ó Gliasáin, M. (1979), *All-Irish Primary Schools in the Dublin Area. Report of a Sociological and Spatial Study of All-Irish-medium Schools in the Greater Dublin Area, with Special Reference to their Impact on Home and Social Network Use of Irish*. Dublin: ITÉ.

Ó Riagáin, P. & Ó Gliasáin, M. (1984), *The Irish Language in the Republic of Ireland 1983: preliminary report of a national survey*. Baile Átha Cliath: Institúid Teangeolaíochta Éireann.

Ó Ríordáin, S. (1964), *Brosna*. Dublin: Sáirséal – Ó Marcaigh.

Ó Siadhail, M. (1989), *Modern Irish: grammatical structure and dialectal variation.* Cambridge: Cambridge University Press.

Ó Torna, C. (2005), *Cruthú na Gaeltachta 1893-1922*. Dublin: Cois Life.

Ó Tuathaigh, G. (1990), *The Development of the Gaeltacht as a Bilingual Entity. Occasional Paper No. 8.* Dublin: Institiúd Teangeolaíochta Éireann. Cited in Ó Riagáin (1992).

Page, R. (2012), "The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags" in *Discourse & Communication*, Vol. 6, No. 2.

Paolillo, J. C. (2001), "Language variation in Internet Relay Chat: A social network approach" in *Journal of Sociolinguistics*, Vol. 5, No. 2.

Paolillo, J. C. (2011), "'Conversational' codeswitching on Usenet and Internet Relay Chat" in Language@Internet, Volume 8. Retrieved from web on 10/10/2013: http://www.languageatinternet.org/articles/2011/paolillo

Pihlaja, S. (2011), "Cops, popes, and garbage collectors: Metaphor and antagonism in an atheist/Christian YouTube video thread" in Language@Internet, Volume 8. Retrieved from web on 10/10/2013: http://www.languageatinternet.org/articles/2011/pihlaja

Pimienta, D., Prado, D. & Blanco, Á. (2009), "Twelve years of measuring linguistic diversity in the Internet: balance and perspectives" [online] in *UNESCO publications for the World Summit on the Information Society*. UNESCO. http://unesdoc.unesco.org/images/0018/001870/187016e.pdf

Preece, J. & Maloney-Krichmar, D. (2003), "Online communities" in Jacko, J. & Sears, A. (Eds.), *Handbook of human-computer interaction*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rheingold, H. (1987), "Virtual communities: exchanging ideas through computer bulletin boards" in *Whole Earth Review*, Winter. Sausalito, California: Stewart Brand.

Rheingold, H. (1993), *The Virtual Community.* Cambridge, Massachusetts: MIT Press.

Robinson, T. (2012), *Connemara: A Little Gaelic Kingdom*. London: Penguin.

Rowe, C. (2011), "Whatchanade? Rapid language change in a private email sibling code" in Language@Internet, Volume 8. Retrieved from web on 10/10/2013: http://www.languageatinternet.org/articles/2011/rowe

Scannell, K. P. (2007), "The Crúbadán project: Corpus building for under-resourced languages". Sourced on 16th February 2012 from: http://borel.slu.edu/pub/wac3.pdf

Scannell, K. P. (2011), "Indigenous Tweets" website: http://www.indigenoustweets.com. Accessed on: 15th Februrary, 2012.

Scannell, K. P. (2013), "Teangacha Mionlaigh sa Ré Dhigiteach: Tionchar na Meán Sóisialta". Presentation to the "Is ar scáth a chéile a mhaireann na daoine" Forum, organised by Fondúireacht na Gaeilge, 2nd March, 2013. Available online at: http://borel.slu.edu/pub/arscath.pdf.

Schuster, M. (2004), "Applying Social Network theory to a small Weblog community: Hubs, Power Laws and the Community Core" in Burg, T. N. (Ed.), *BlogTalks 2.* Norderstedt.

Scott, J. (2000), *Social Network Analysis, a handbook (2nd edition)*. London: Sage Publications.

Severinson Eklundh, K. (2010), "To quote or not to quote: Setting the context for computer-mediated dialogues" in Language@Internet, Volume 7. Retrieved from web on 10/10/2013: http://www.languageatinternet.org/articles/2010/2665

Siebenhaar, B (2006), "Code choice and code-switching in Swiss-German Internet Relay Chat rooms" in *Journal of Sociolinguistics*, Vol 10, No. 4. Oxford: Blackwell.

Smith, M. (1992), "Voices from the WELL: The logic of the virtual commons", Masters Thesis, UCLA. Available online at: http://www.sscnet.ucla.edu/soc/csoc/papers/voices/.

Soanes, C. & Stevenson, A. (2005), *Oxford English Dictionary (2nd edition, revised)*. Oxford: Oxford University Press.

Snyder, J. (1996), "Get real" in *Internet World*, Vol. 7, No. 2. Westport, CT: Mecklermedia.

Stone, A. R. (Sandy) (1991), "Will the real body please stand up? Boundary stories about virtual cultures" in Benedikt, M. (Ed.), *Cyberspace: First steps*. Cambridge, MA: MIT Press.

Stratton, J. (1997), "Cyberspace and the globalization of culture" in Porter, D. (Ed.), *Internet culture*. New York: Routledge.

Suttles, G. (1972), *The social construction of communities*. Chicago: University of Chicago Press.

Thurlow, C., Lengel, L., Tomic, A. (2004) *Computer Mediated Communication: Social Interaction and the Internet*. London: SAGE Publications.

Trammell, K. D., Tarkowski, A., Hofmokl, J. & Sapp, A. M. (2006), "Rzeczpospolita blogów [Republic of Blog]: Examining Polish bloggers through content analysis" in *Journal of Computer-Mediated Communication*, Vol. 11, No. 3.

Turkle, S. (1995), *Life on the Screen: Identity in the age of the Internet*. New York: Touchstone.

Turkle, S. (2010a), *Alone Together: Why we expect more from technology and less from each other*. New York: Basic Books.

Turkle, S. (2010b), "Social media make us less human, warns US expert as backlash mounts" in *The Observer*, 23/01/2010.

Twitter Blog entry (2012, March 21), "Twitter turns six". Retrieved from http://blog.twitter.com/2012/03/twitter-turns-six.html

Uncapher, W. (2007), "Electronic Homesteading on the Rural Frontier: Big Sky Telegraph and its Community" in Bell, D. & Kennedy, B. (Eds.) *The Cybercultures Reader* (2nd Edition). Abingdon, Oxon: Routledge.

Wagner, H. (1958), *Linguistic Atlas and Survey of Irish Dialects, Vol. I, Introduction, 300 maps*. Dublin: Dublin Institute for Advanced Studies.

Wardaugh, R. (2002), *An Introduction to Sociolinguistics, Fourth Edition*. Malden, MA: Blackwell Publishing.

Warschauer, M. (2002), "Language.com: The Internet and linguistic pluralism" in Snyder, I. (Ed.) *Silicon literacies: Communication, innovation and education in the electronic age*. London: Routledge.

Wasserman, S. & Faust, K. (1994), *Social Network Analysis, Methods and Applications*. Cambridge: Cambridge University Press.

Wasserman, S., Scott, J. & Carrington, P. J. (2005) "Introduction" in Carrington, P. J., Scott, J. & Wasserman, S. (Eds.), *Models and Methods in Social Network Analysis*. Cambridge: Cambridge University Press.

Webber, M. M. (1963), "Order in diversity: community without propinquity', in Wingo, L. (Ed.), *Cities and space.* Baltimore: Johns Hopkins University Press.

Wellman, B. (2001). Message posted to AIR-L@, December 25, 2001.

Wellman, B. (2002), "Little Boxes, Glocalization, and Networked Individualism" in Tanabe, M., van den Besselaar, P. & Ishida, T. (Eds.), *Digital Cities II: Computational and Sociological Approaches*. Berlin: Springer.

Welser, H. T., Gleave, E., Fischer, D. & Smith, M. (2007), "Visualizing the signatures of social roles in online discussion groups" in *The Journal of Social Structure*, Vol. 8, No. 2.

Wilson, M. (2007), "Community in the Abstract: A political and ethical dilemma" in Bell, D. & Kennedy, B. (Eds.) *The Cybercultures Reader* (2nd Edition). Abingdon, Oxon: Routledge.

Wodak, R. & Wright, S. (2006), "The European Union in cyberspace: Multilingual democratic participation in a virtual public sphere?" in *Journal of Language and Politics*, Vol. 5, No. 2.

Wolfram, W. (1997), 'Dialect in Society' in Coulmas, F. (Ed.), *The Handbook of Sociolinguistics*. Oxford: Blackwell Publishing.

Wright, S. (2006), "Regional or Minority Languages on the WWW" in *Journal of Language and Politics*, Vol. 5, No. 2.