

VILNIUS UNIVERSITY
FACULTY OF PHILOLOGY

Benas Venckaitis

Programme of English Studies

Clitics in Celtic Languages in Typological Perspective

MA thesis

Supervisor: Dr. Doc. Vladimir Panov

VILNIUS
2024

Table of contents

Abstract	4
1. Introduction	5
1.1 A brief history of linguistic typology.....	5
1.2 The clitic phenomenon.....	8
1.3 Clitic or affix?.....	10
1.4 The canonical clitic versus a non-aprioristic approach.....	11
1.5 The heart of the clitic problem.....	14
1.6 Clitics in European languages.....	16
2. Data and methods	17
3. Results and discussion	18
3.1 Introduction to Celtic languages.....	18
3.2 Goidelic languages.....	19
3.2.1 Irish.....	19
3.2.1.1 An overview of Irish.....	19
3.2.1.2 A historicity of Irish.....	19
3.2.1.3 Clitics in Modern Irish.....	20
3.2.2 Scottish Gaelic.....	24
3.2.2.1 A historicity and overview of Scottish Gaelic.....	24
3.2.2.2 Clitics in Scottish Gaelic.....	24
3.2.3 Manx.....	27
3.2.3.1 A historicity and overview of Manx.....	27
3.2.3.2 Clitics in Manx.....	27
3.2.4 A summary of results in Goidelic languages.....	30
3.3 Brythonic languages.....	32
3.3.1 Welsh.....	32
3.3.1.1 An overview of Welsh.....	32
3.3.1.2 A historicity of Welsh.....	32
3.3.1.1 Clitics in Modern Welsh.....	33
3.3.2 Breton.....	36
3.3.2.1 An overview of Breton.....	36
3.3.2.2 A historicity of Breton.....	37

3.3.2.3 Clitics in Modern Breton.....	38
3.3.3 Cornish.....	40
3.3.3.1 An overview of Cornish.....	40
3.3.3.2 A historicity of Cornish.....	41
3.3.3.3 Clitics in Cornish	41
3.3.4 A summary of Brythonic languages results.....	43
4. Conclusion.....	45
5. Data sources.....	47
6. References	49
7. Summary in Lithuanian	51

Abstract

This paper delves into the linguistic phenomenon known as ‘clitics’. By its most simple and widely accepted definition, clitics are short linguistic forms that have the syntactic characteristics of a word, but phonologically depend on another word or phrase. However, researchers often disagree on the exact definition of the phenomenon, or occasionally even its existence. This paper will provide a brief overview of the history of research conducted about this term (Haspelmath 2015, 2023; Spencer & Luis 2012; Zwicky 1977; Zwicky & Pullum 1983) as well as attempt to narrow it all down to a definition applicable in typological analysis. Said definition will then be used to identify clitics in the Celtic language family, namely the Irish, Scottish Gaelic, Welsh, Manx, Breton and Cornish languages. Previous research has shown that these languages exhibit the aforementioned phenomenon to varying extents, however, little cross-language comparison between them has been conducted so far. The results will then be analyzed and a typology of possible clitic particles in Celtic languages drawn. Conclusions regarding the usage of this linguistic form will be made, as well implications for further research in linguistic typology on the elusive, and at times controversial, subject of clitics.

keywords: clitics; Celtic languages; linguistic typology; morphology; syntax; comparative concept

1. Introduction

1.1 A brief history of linguistic typology

The human need to categorize everything is amply present in many aspects of society: for example, the buildings one sees in the street will belong to one or another architectural category, the cars one can drive range from a hatchback to a pickup truck, or the ice cream that one eats on a hot summer day may be of many types, such as gelato or sorbet. Even humans themselves are classified by gender, sex, race, age, nationality, etc. “As humans, we seem to have a penchant for classifying and labelling people and things. It helps us establish a sense of order in the world around us” (Miller 2020: 439). The need for order in life is a natural part of being, as it helps us to understand the surrounding world. As such, it only seems natural that much like everything else that is part of the human existence, language should be classified as well. This idea is where the field of linguistic typology gets its origins.

The notion of classifying languages in order to compare them and draw implications about other languages is not entirely new and predates the field of linguistic typology itself. As early as the Middle Ages scholars began to notice the existence of different linguistic systems, comparing various European languages with Latin, with it being “the basis for a general theory of grammar” (Ramat 2010: 2). In his *De vulgari eloquentia* (around 1303), the poet, writer and philosopher Dante Alighieri took a big step towards what is considered modern language classification by making a distinction between three language families in Europe – Germanic, Latin and Greek. As Renaissance came about in the coming centuries, these speculations about language classification only increased. One of the more notable works on the subject is Jean de Drosay’s *Grammaticae quadrilinguis partitiones* (1544), which is considered to be the first comparative grammar, comparing the grammars of French and the three ‘holy’ languages of the time – Hebrew, Greek and Latin. While this was still far from what is considered linguistic typology, the fact remains that the Renaissance period saw the first proper cross-linguistic comparison (Ramat 2010).

As scholars began to be exposed to more and more new languages with the discovery of the New World, ‘missionary linguists’ began to appear, charting the linguistic map with a bouquet of the novel, exotic languages. This merited a categorisation and comparison of some sort, and while adapting the descriptions of New World languages to the grammatical patterns of Western tradition was not easy, a notable work was published in 1800 by the Spanish Jesuit and philologist Lorenzo Hervás y

Panduro by the name of *Catálogo de las lenguas de las naciones conocidas*, which, as the name implies, provided a catalogue of the languages known at the time to the Western scholarly world. Similar works of language cataloguing and categorisation were continued to be published around that time, such as Johann Cristoph Adelung's 1817-published language sample with a wordy title *Mithridates, oder allgemeine Sprachen-kunde mit dem Vater Unser als Sprachprobe in nahe fünfhundert Sprachen und Mundarten*. "The underlying idea was that the cross-linguistic comparison of very different languages may be able to uncover the general philosophical principles – that is, the *characteristica universalis* – and, at the same time, recover the evolution of man's faculty of language" (Ramat 2010: 4). Around the same time, in 1816, Franz Bopp wrote his famous 'Conjugation system' (with a wordy title in German as well), a book considered to be a milestone in the beginning of modern linguistics. In it, Bopp does ground laying work for historical linguistic comparison, positing a common origin for languages that are alike – effectively starting comparative Indo-European language studies as an academic discipline. Other authors have had their input in this field as well, such as August Wilhelm Schlegel, whose 1818 work *Observations sur la langue et la littérature provençales* "led to a first typological division of languages" (Ramat 2010: 5). Schlegel divided languages into three classes – a) without grammatical structure, b) agglutinative, and c) inflectional, which can be analytic or synthetic. These ideas were later expanded upon by his contemporaries August Schleicher, Heymann Steinthal, Franz Misteli, Franz Nicolaus Finck, and Max Müller (Ramat 2010). One work of literature after another, the foundations for the field of linguistic typology were being laid.

The first step towards what is now considered to be linguistic typology is believed to have been taken by Georg von der Gabelentz in 1901 when he christened the field, giving it the name 'typology' (Jung Song 2010). Of course, he did more than just that – the scholar "argued that classification of languages on a genealogical basis was not to be equated with classification of languages based on linguistic types", as well as pointing out that "there was no scale according to which languages could be more perfect than others, innovatively stepping away from the evaluative notions that had previously been associated with language types" (Velupillai 2012: 1). It was this approach that formed the field and made it stand out from the previously found bias when classifying languages, leaving behind the notion that some language types are superior to others. Von der Gabelentz also indicated a need to search for a "unique, overarching principle" and a "possible all-embracing explanation" – in other words, that modern typology should go beyond simple classification and categorization and aim to become predictive (Ramat 2010). In the years following, other linguists further developed these ideas. Edward Sapir stated that an "adequate classification of languages must be based on the nature of

the concepts expressed by the language”, essentially popularizing the statements made by von der Gabelentz (Graffi 2010: 7). Louis Hjelmslev posited that typology cannot be based only on the analysis of the word: “<...> it must first of all investigate the cross-linguistic relationships between structural categories (such as accent, pitch, etc., on the ‘expression plane’, and case, gender, number, etc., on the ‘content plane’) and then those between the categories of usage” (Graffi 2010: 8). This newborn area of linguistics still had a way to go before its scholars could properly agree on a framework; it would take a few more decades before that would happen.

The relative chaos of linguistic typology was famously contained in the 1960s by the work of American linguist Joseph Greenberg, for example, his 1963 paper on a cross-linguistic survey of word order which would lead to a number of implicational universals. Greenberg’s work is closely associated with the comparative tradition of Roman Jakobson and the Prague School, a language and literature society which gave emphasis to a structuralist linguistic analysis of languages as systems of functional units (Vellupillai 2012). This began a new period of linguistic typology where its focus shifted from morphology to morphosyntax, with Greenbergian theory serving as a model for the field (Jung Song 2010). Potential universals and cross-linguistic tendencies as well as their implications became the primary object of study. The later decades saw typologists taking note of non-linguistic factors at work as well, such as language contact, number of speakers, movement of population, etc. (Jung Song 2010). A notion came about that absolute language universals are a fallacy, given the limited current state of knowledge of all the languages in the world and their grammars. More new questions arose in the new millennium. Typologists “have come to the realization that many typological properties are not evenly distributed in the world and have begun to ask in earnest why” (Jung Song 2010: 2). As such, at the current moment linguistic typology is a field on the rise, constantly positing new issues and adopting new methods and approaches to tackle them. With the advent of advanced technology, large-scale databases are made available to typologists, which serves to advance studies in the field. The future looks promising as well – with new empirical findings and problems of explaining why linguistic diversity is the way that it is, it does not seem that the field will run out of topics to discuss any time soon (Jung Song 2010).

Having briefly run through the history of linguistic typology, it would be useful to provide a single, comprehensive definition of what it entails. However, different scholars provide slightly different accounts. Michael Daniel (2010:1) claims that “Linguistic typology compares languages to learn how different languages are, to see how far these differences may go, and to find out what generalizations can be made regarding cross-linguistic variation”. Paolo Ramat (2010: 1) writes that it

is “the systematic cross-comparison that aims to discover the underlying universal properties of human language.” Viveka Velupillai (2012: 1) provides a brief description of it being “the systematic study and comparison of language structures”. It seems the most exhaustive (and one of the most recent) definitions is provided by Jae Jung Song (2018), wherein he provides many examples of how typology can be applied to find out the nature of human language. Perhaps that is the answer: linguistic typology is a comparison of languages and their features with the explicit goal of making implications from the results of these comparisons.

1.2 The clitic phenomenon

Within the framework of linguistic typology there are many interesting research problems not counting those already mentioned above, such as defining the general categories of natural languages or the problem of language universals (Ramat 1987). Perhaps one of the most interesting ‘typological puzzles’ that still merits a lot of research is the phenomenon of clitics. To give a very brief and inexhaustive description of a clitic provided by the SIL Glossary of Linguistic Terms: “A clitic is a morpheme that has syntactic characteristics of a word but shows evidence of being phonologically bound to another word” (SIL Glossary of Linguistic Terms, n.d.). In other words, a clitic may appear like an affix, but what makes it different is its syntactic role at the phrase level – it might have the form of an affix, but it has the distribution of a function word. This definition might appear clear enough at first glance. However, the problem is that there are so many morphemes that fall out of this definition and are still called clitics, that one might call this definition the exception rather than the rule. Linguists have not yet agreed on a single set of properties that always distinguishes what clitics actually are, which makes this phenomenon one of the more interesting ‘typological puzzles’ that the field has to offer.

An account of the various approaches to the clitic problem is in order. There have been works defining the existence of clitics as early as the 19th century, for example, Wackernagel’s 1892 paper *On a law of Indo-European word order*, which takes second-position clitics in Indo-European languages as its focus. However, it was only in 1977 that the first major attempt to properly classify clitics within a genuinely cross-linguistic perspective was published by Zwicky (Spencer & Luís 2012). The scholar introduced a new clitic typology which made a bipartite classification between simple and special clitics. As the name implies, simple clitics feature in regular syntax, while special clitics exhibit an idiosyncratic position in syntax. One of Zwicky’s key features of a simple clitic is weak phonology, it

being “phonologically reduced, <...> phonologically subordinate to a neighboring word” (Zwicky 1977: 5). It also exhibits regular syntax – the syntactical position is identical to its free (stressed) counterpart. Spencer and Luís (2012) indicate English personal pronouns and prepositions as examples of simple clitics, as shown in (1a) and (1b):

- (1) a. *She met him*
Full form: [mɛt hɪm]
Clitic form: [mɛtɪm]
- b. *Mary looked at me*
Full form: [æt]
Clitic form: [ət]

By contrast, special clitics behave differently, not necessarily appearing in the same place as its full form syntactically, only acting as a “variant of a stressed free form with the same cognitive meaning and with a similar phonological makeup” (Zwicky 1977:3). The scholar gives weak or clitic pronouns of many Romance and Slavic languages as examples, for instance, how in French declarative sentences conjunct object pronouns are placed before the verb despite French declarative sentence being SVO. (2a) shows a standard French SVO sentence, whereas (2b) depicts a contrasting word order with a clitic:

- (2) a. *Je vois Jean* [ʒə vwa ʒɑ̃]
‘I see John’
- b. *Je le vois* [ʒə lə vwa]
‘I see him’

Other Romance languages feature similar examples. According to this notion, what makes clitics special is their unusual syntactic placement pattern. Zwicky’s typology has been highly influential as a starting point of clitic classification; however, it has since been subject to criticism. In his paper, the author also denoted a third class of clitics – the bound word, defining it as an unaccented form which has no accented counterpart, for example, the English possessive ‘s (Zwicky 1977). But as Spencer and Luís point out, “a clitic can sometimes be very different from its full-form counterpart, both in phonological form and in morphosyntactic behavior, and then it becomes very difficult to know

whether we can really say that a clitic has a full form” (2012: 43). As such, distinguishing between special clitics and bound words may become another challenge, which was addressed by Zwicky himself in his 1985 paper wherein he subsumed bound words under the class of special clitics. There are more problems with this typology – for example, as Halpern (1998) notes, simple clitics may not always have the same distribution as their full-form counterparts, instead defining them as positioned in a subset of the positions in which the full forms are found, further broadening the scope of their classification. Also, Anderson (2005) points out that clitics being incapable of bearing stress or other forms of prominence is not always true as well. As such, he proposes classifying simple clitics as ones of phonological weakness, and special clitic as ones of unusual syntax.

1.3 Clitic or affix?

Another problem arises when attempting to establish criteria for identifying clitics – the distinction between a clitic and an affix. Taking into account the definitions of clitics already discussed above, it is clear that affixes may in some respects fall under the proposed category of bound words, but that cannot be the case – a morpheme cannot be both an affix and a clitic. This is where Zwicky and Pullum come in with their arguably most influential paper on this topic *Cliticization vs. inflection: English n't* (1983). In it they present a set of criteria for clitichood and affixhood, commonly known as the Zwicky-Pullum criteria:

- 1) Host selectivity: Clitics can exhibit a low degree of selection with respect to their hosts while affixes exhibit a high degree of selection with respect to their stems.
- 2) Arbitrary gaps in the set of combinations are more characteristic of affixed words than of clitic groups.
- 3) Morphophonological idiosyncrasies are more characteristic of affixed words than of clitic groups.
- 4) Semantic idiosyncrasies are more characteristic of affixed words than of clitic groups.
- 5) Lexical integrity: Syntactic rules can affect words, but cannot affect clitic groups.
- 6) Clitic-affix ordering: Clitics can attach to material already containing

clitics, but affixes cannot.

These criteria have been widely attested as a good way to distinguish clitics from affixes, but yet again, as Spencer and Luís point out, “there are plenty of instances of things that are traditionally called clitics violating one or other of these criteria <...> These criteria therefore indicate tendencies and are not defining characteristics that allow us to determine with absolute certainty whether a given formative is an affix or a clitic” (2012: 110). Zwicky and Pullum address this by discussing the English formative *-n't*, which is usually referred to as a clitic, however, as the researchers find, as a clitic it only matches one of the six aforementioned criteria, reaching a conclusion that it must be an affix. Spencer and Luís discuss this by providing more similar examples of clitics and their clusters with more affix-like properties, such as Italian or Luiseño, and at one point even conclude that is “more or less impossible to propose a model of grammar in which there is a lexicon of clitics that remain clitics in in all respects”, suggesting that one “shouldn’t actually draw a clear distinction between clitics and affixes <...> the obvious conclusion is that special clitics are in fact special affixes, special in the sense that their placement is defined in terms of syntactic and prosodic constituents and not just in terms of classes of stems” (2012: 218). From this, one could infer that clitics and affixes are not in fact hard opposites, but rather exist on a scale, with simple clitics and normal affixes being on both ends, while special clitics and special affixes intersect somewhere in the middle.

1.4 The canonical clitic versus a non-aprioristic approach

In an effort to combat this linguistic beast, Spencer and Luís wrote the article *The canonical clitic* (2012). The notion of a grammatical concept being ‘canonical’ comes from the linguistic subfield of Canonical Typology. The key idea of this field of research is establishing canonical criteria upon which to determine a ‘canon’ or ‘canonical ideal’ – “a reference point from which to compare linguistic objects and descriptions” (Bond 2018: 2). These canonical criteria are established by first defining a base domain of features associated with the phenomenon in question that allow a linguistic structure to be considered an example of the phenomenon. Then out of the base parameters those which are considered to be “the best, the clearest, the indisputable ones” (Corbett 2006: 141), in other words, those which seem the most logically determined, are selected to be the canonical properties of a phenomenon. Out of these criteria a logically motivated canon is then defined as a “theoretical reference point from which to calibrate the sample population” (Bond 2018: 24). Phenomena may then

be described as either canonical (adhering to these criteria) or non-canonical (failing to meet one or several criteria but still considered to be a correct example).

First describing the typical, or canonical, characteristics of what a word is and what an affix is, the authors reach this definition of the canonical clitic: “canonical clitic has the canonical form properties of an affix and the canonical distributional properties of a function word. In other words, the canonical clitic lies at the ‘intersection’ of the space defined by the dimensions ‘affix form’ and ‘function word distribution’ (Spencer & Luís 2012: 149). Some argumentations for setting this set of criteria could be that most linguistic items have some sort of a canonical definition based upon which further subsets of classes or deviations from the norm can be identified. As such, when discussing clitics, it is simply rudimentary that a certain hard-set definition be established, before even beginning to discuss their various forms and deviations. Upon attempting to identify whether a morpheme is a clitic or an affix, the canonical clitic should be the first point of analysis. Of course, this is by no means an ‘applicable-to-all’ clear-cut distinction, but it seems that Spencer and Luís believe Canonical Typology should accept this definition as a starting point.

However, not all linguists have been jumping at the chance to accept Spencer and Luís’ canonical clitic approach as the fundamental one when it comes to discussing and identifying clitics. Haspelmath (2015) presents some different ideas in his article *Defining vs. diagnosing linguistic categories: A case study of clitic phenomena*. In it the researcher discusses the idea of the canonical clitic as presented by Spencer and Luís (2012), claiming that “the great majority of clitic phenomena that have been discussed in the relevant literature (e. g., Spencer and Luís 2012) are not canonical clitics <...> some elements which nobody would call a clitic end up as canonical clitics” (2015: 290). He goes on to argue that canonical definitions are simply not applicable on large scale cross-linguistic research because of disparateness between languages and non-convergence between identifying properties of clitics. What Haspelmath offers instead is a non-aprioristic approach, essentially stating that clitics in different languages should be looked at objectively, in a sort of case-by-case scenario. In order to understand this better, it is imperative to reject the notion of pre-established linguistic categories. Haspelmath (2007) states that instead of using formal categories, one must consider semantic-pragmatic or phonetic substance when making classifications and generalizations. For example, it cannot be presupposed that for a typology of passive constructions, every language conveniently possesses a concept of “passiveness” with diverse modes of expression that can be studied. But what typologists can do is find some level of meaning at which the passiveness of languages can be compared. If one can translate a low-level notion like ‘to be done’ into another

language, a semantic relation can be defined between ‘to’, ‘be’, ‘done’ as “passive”. “As long as there is translatability of simple concepts, comparison should be possible” (Haspelmath 2007: 8).

Haspelmath expands on this in a later paper discussing clitics wherein he introduces two new terms, plenimorphs (3a) and minimorphs (3b):

- (3) a. plenimorph = a morph that denotes a thing, a process, or a property (= a root)
- b. minimorph = a morph with a meaning that is normally omitted in translation into some other language without significant loss of content (i.e., a meaning that could easily be inferred from context, or a meaning that makes a small, subtle contribution)

Basically, according to the author, minimorphs include “articles, case forms such as accusative or genitive, bound person markers, tense forms such as future or past, complementizers, and coordinators like ‘and’”, while plenimorphs (or roots) are the main bases of word-meaning which are usually kept when translated across languages, as opposed to minimorphs (2015: 293). Furthermore, as minimorphs coalesce with other morphs, they do so variably – so, say, if a minimorph is coalescent on more than 80% of the coalescence properties (prosodic dependency, adjacency, narrow scope, shape idiosyncrasy), it is probably an affix. But if a minimorph is weakly coalescent, verging on 30-50% of the coalescence properties, it might be attached the label of a clitic. Haspelmath does not purport this to be a universal taxonomy, but instead offers it simply as an alternative to the restrictive approach of the canonical clitic.

Furthermore, in his 2023 paper *Types of clitics in the world’s languages* the scholar further refutes Spencer and Luís’ proposed definition by stating that clitics are not in fact intermediate between words and affixes – instead, he defines clitics as “a comparative concept”, rather than being a fixed grammatical concept. The term was introduced by Haspelmath earlier in earlier works, and he describes it as a concept “created by comparative linguists for the specific purpose of crosslinguistic comparison” (2010: 665). Comparative concepts are designed to be universally applicable, unlike traditional linguistic categories. In this view, “the fact that two language-particular categories both match a comparative concept just means that they are similar in the relevant respect, but not that they are ‘the same’ in any sense” (Haspelmath 2010: 666). Once a comparative concept, for example, a relative clause, has been established, then a language-particular descriptive category can be made to define a linguistic construction which might be given the label of a relative clause. For instance, Japanese has no category that matches the descriptive category of the relative clause in English, but it’s language-

specific category of ‘Noun-Modifying Construction’ may be equated to the comparative concept of the relative clause in the comparative sense (Haspelmath 2010). As such, it might be considered that a clitic is not a clear-cut linguistic category, but simply a comparative concept created for easier language comparison.

Another interesting statement by the German linguist asserts that “just as “clitic” is not more than a comparative concept with some usefulness for linguists, the familiar affix vs. word distinction (and the morphology vs. syntax division in grammar) could be largely based on the orthographic word” (Haspelmath 2023: 33). Indeed, when attempting to identify a clitic, there is a need not only to distinguish it from an affix, but also from a full-fledged free word (if there is in fact a clear, discernable difference). And here another typological problem arises – just how exactly does one define a word?

1.5 The heart of the clitic problem

What may look like a rather simple question to answer for linguists is actually more complex than at first glance would tell. As it happens, linguists have not been able to agree on a single clear and exhaustive definition of what a word is, a seemingly comical situation which has even been called a crisis of the field. As Haspelmath himself writes in his paper *The indeterminacy of word segmentation and the nature of morphology and syntax*, “Linguists generally employ a range of different criteria, but these are not uniformly applicable across contexts and languages, and where they are applicable, they do not always converge” (2011: 2). Nonetheless, despite the different criteria, there appears to be some agreement regarding the categories of words. Dixon and Aikhenvald (2003) discern these types:

1. **Grammatical word**, consisting of a number of grammatical elements which:
 - (a) always occur together, rather than scattered through the clause (the criterion of cohesiveness);
 - (b) occur in a fixed order;
 - (c) have a conventionalised coherence and meaning.
2. **Phonological word** - a phonological unit larger than the syllable (in some languages it may minimally be just one syllable) which has at least one (and generally more than one) phonological defining property chosen from the following areas:
 - (a) Segmental features – internal syllabic and segmental structure; phonetic realisations in terms of this; word boundary phenomena; pause phenomena.

(b) Prosodic features – stress (or accent) and/or tone assignment; prosodic features such as nasalization, retroflexion, vowel harmony.

(c) Phonological rules – some rules apply only within a phonological word; others (external sandhi rules) apply specifically across a phonological word boundary.

3. **Orthographic word** - used to refer to a single word in writing that is separated by spaces on either side.

It is worth noting that these categories are not exclusive to each other and often coincide. “In English, for instance, there are examples of two grammatical words making up one phonological word, e.g. *don't*, *won't*, *he'll*” (Dixon & Aikhenvald 2003: 24). Interestingly enough, this would imply that *-n't*, which is by many considered to be a clitic, is to be categorized as a grammatical word. The 1983 Zwicky and Pullum paper on *-n't*, discussed above, would seem to disagree, and classify it as an affix-like special clitic.

This is where the heart of the clitic problem lies. Since there is no clear universally applicable definition of ‘word’, it is not easy to place clitics, words, and affixes apart. We have seen clitics that behave word-like, and sometimes form full orthographic words. They can be called grammatical words – what Zwicky called ‘simple clitics’. These do not form a full phonological word, as they must attach to a host – a morpheme with which it forms one phonological word. Words, but not really words. We have also described what Zwicky called ‘special clitics’, ones that behave like affixes but are syntactically separate. Again, phonological dependency is implied. Affixes, but not really affixes. The main problem is applying these criteria across languages. Affixes and especially words differ in their definitions in different language families. Aikhenvald (2002) advocates for a continuum between free words, clitics, and affixes, but Haspelmath states that there “is no reason to think there is a single scale or continuum” (2023: 33). As such, it seems that one really must go for a non-aprioristic case-by-case, language-by-language scenario when attempting to identify this linguistic phenomenon.

The most recent definition of a clitic is offered by Haspelmath: “A clitic is a bound morph that is neither an affix nor a root” (2023: 19). The author claims that this is the superior definition to consider because it signifies two most important aspects to clitic identification – a) the lack of word-class selectivity, or indiscriminacy, which separates it from an affix, and b) not having a root, which separates it from simple nouns, adjectives, and verbs. Haspelmath also dismisses the above discussed phonological criterion, as “this criterion would be significant only if we already knew that clitics are words. If we could initially divide the types of (non-phrasal) forms into affixes and words, <...> then

the prosodic deficiency of clitics would be relevant” (2023: 20). Once again as before, the researcher states that this definition is not necessarily universally applicable, but that it does help to think of clitics as a comparative concept and compare them in different languages. Therefore, this paper will primarily make use of this particular definition, while keeping in mind other approaches to clitic defining as well.

1.6 Clitics in European languages

A great many languages have particles and morphemes in them that have been described as clitics by various researchers. Taking into account Indo-European languages alone, there are a lot of articles describing the clitic systems of such variant-rich language families as Slavic, Romance and Germanic, wherein the discussion about these particles is plentiful (Franks 2000). However, not all European language families enjoy such a vast description and categorization of their prevalent clitic phenomena. One such example is the Celtic language family. Although there are some descriptions of clitic particles found in Celtic languages, they are somewhat scattered and inconsistent.

In order to possess a better understanding of cross-linguistic trends and the way that different phenomena function across different languages, a comprehensive account of Celtic linguistic structures that may be called clitics is needed. The aim of this paper is precisely that – a collection, analysis and categorization of Celtic clitic particles that will allow more adequate insight into the linguistic structures of this language family. Not only that, but the results of this study may help to contextualize Celtic clitic particles within a broader Indo-European context. The study also contributes to the study of Europeanisms and SAE (Standard Average European) by adding new phenomena descriptions to the field, and the results of it may be used to enrich the field of areal linguistics.

2. Data and methods

To perform a cross-linguistic analysis of Celtic languages, two types of sources had to be taken – books with descriptive grammars of the languages and published articles dealing with some grammatical aspect of the languages. The total sources analyzed are comprised of 3 books overviewing Celtic grammar in general, 9 books detailing the particular grammars of the Celtic languages (1 of Breton, 1 of Cornish, 2 of Irish, 1 of Manx, 1 of Old Irish, 1 of Scottish Gaelic, 2 of Welsh), and 31 published articles focusing on a relevant grammatical aspect (clitics, syntax, morphology, agreement, etc.) of the particular language (9 of Breton, 1 of Cornish, 6 of Irish, 1 of Manx, 4 of Old Irish, 2 of Scottish Gaelic, 7 of Welsh, and 1 of general Celtic syntax). Note that only the currently living Celtic languages were analyzed.

The grammars and articles were read and scoured for particles that could be considered as clitics. To do so, Haspelmath's most recent (2023) definition of what a clitic may be (neither an affix nor a root) was applied, although other definitions were occasionally applied as well. In some sources there were already categories identified as clitics by the authors – in these cases, the aforementioned clitic definition was applied as well to the proposed categories to see if they match Haspelmath's comparative concept. In other cases, possible clitics were identified manually. The examples picked out were assigned categories (for example, the determiner clitic) and these categories then compared cross-linguistically. First, the clitics found in Goidelic (Irish, Scottish Gaelic, Manx) sub-family group languages were analyzed together, then the ones found in the Brythonic (Welsh, Cornish, Breton) languages faced the same treatment as well. Finally, all clitic categories across the Celtic languages were cross-linguistically compared and conclusions based on their identification and usage made, with suggestions for future research and use of results posited.

3. Results and discussion

3.1 Introduction to Celtic languages

As different as languages may be, from a typological standpoint it still makes more sense to group them together than to keep them apart. There are several such ways of language grouping, and one of the most popular ways is to put them into language families based on similar structural features. These groups are typically comprised on the basis of having a shared ancestral, or proto-language, as well as various common features typical of or unique to the group. One such group is called Celtic, an Indo-European language family that had once spread across the whole of central Europe, even up to the region of Anatolia in modern day Turkey. Today, however, surviving Celtic languages are confined to northwestern Europe, mainly the British Isles. They are considered to be of the Insular Celtic subdivision, as their main period of development took place on the islands mentioned. There is also the Continental Celtic subdivision, comprised of languages recognized to be extinct at this point, such as Gaulish or Galatian (the one spoken in Anatolia). The completeness of attestation between these languages is highly varied, as not much written evidence has been left behind for researchers to analyze (Fife 2009). As such, they will not be the subject of this paper.

As stated by Fife, “<...> meaningful discussion of the typology of Celtic requires one to confine attention primarily to the so-called neo-Celtic languages, the languages attested in the post-Roman era” (2009: 6). Member of the Insular Celtic group are divided further as follows:

- ***Gaelic, or Goidelic:***
 - a) **Irish.** Spoken in the Republic of Ireland. Speakers: approx. 1.8 million (native: approx. 70,000)
 - b) **Scottish Gaelic.** Spoken in Scotland, United Kingdom. Speakers: approx. 60,000
 - c) **Manx.** Spoken in the Isle of Man, United Kingdom. Speakers: approx. 2,000 (native: 23)

- ***British, or Brittonic:***
 - a) **Welsh.** Spoken in Wales, United Kingdom. Speakers: approx. 530,000
 - b) **Breton.** Spoken in Brittany, France. Speakers: approx. 210,000
 - c) **Cornish.** Spoken in Cornwall, England. Speakers: approx. 500

Naturally, owing to the number of capable speakers, the number of sources on these languages varies as well – for example, it is much easier to draw a typology of Welsh than it is of Cornish. Nonetheless, the typicality of what relates to linguistic ‘Celticity’ (beautiful term provided by Fife) is quite interesting.

In fact, one of the issues particular to Celtic typology is the phenomena of clitics. It has been attested as a feature of these languages, although more prominent in some than others. As Borsley and Roberts state, “Clitics and, arguably, clitic doubling are pervasive phenomena in the Celtic languages” (1996: 16). So, what exactly is the role of the clitic in Celtic languages and how does it relate to what is already known about clitics?

3.2 Goidelic languages

The currently spoken Goidelic languages include Irish, Scottish Gaelic, and Manx. Below a brief overview, historicity and possible clitic particles will be presented for each language. Afterwards, a summary and analysis of particles found in all three will follow.

3.2.1 Irish

3.2.1.1 An overview of Irish

The Irish language is in a rather unique position from a sociolinguistic perspective. It is recognized as the first official language in the Republic of Ireland according to its Constitution, but this status is much more symbolic than practical. English, denoted as the “second official language”, is in fact the one that has the most widespread daily usage. Irish as a daily language is used only in areas known as ‘Gaeltachtaí’, which are mostly spread out on the western coast of the country and comprise 2% of the total population. Most governmental institutions conduct their business in English, as well as virtually all private businesses (save for those that operate in the Gaeltachtaí) (Ó hIfearnáin 2009). Irish is a compulsory subject in primary and secondary school, which led to the fact that the number of capable speakers according to the 2022 census is around 1.8 million, however, according to the same census, actual daily users comprise approximately 70,000 speakers.

3.2.1.2 A historicity of Irish

Irish is typically divided into these historical periods: Primitive Irish (3rd-6th century); Old Irish (6th-10th century); Middle Irish (10th-13th century); Early Modern Irish (13th-18th century); Modern Irish (18th century-present). What sets the historicity of Irish apart from other Celtic languages is the fact that going back as far as Old Irish the language is attested in far greater amounts than other Celtic languages of similar periods; as such, its large corpus allows for a greater analysis (Stifter 2009).

Particles that are considered clitics by different scholars can be found all across recorded Irish – for example, Old Irish is noted by Stifter (2009) to have a number of demonstrative particles (proximate: *-sa*; *-so*; *-se*, *-sea*, *-seo*; distant: *-sin*) that feature enclitization:

a) *in cú-so*

the dog=this

‘this dog’ (Stifter 2009: 81)

b) *inna ingine-sin*

the girl=that

‘that girl’ (Stifter 2009: 81)

Of the proximate demonstrative particles only *-seo* survived into Modern Irish. Another clitic-candidate morpheme that did not stand the test of time was the deictic particle *í*, typically placed in front of another clitic particle to convey emphasis:

c) *ind í-si*

the DEIC=3SG.F

‘**this** one (fem.)’ (Stifter 2009: 81)

Some clitic particles that have been existent since Old Irish did in fact survive into modern times, namely the pronominal and possessive forms. These are very similar to those found in Modern Irish and, as such, will be analyzed in the following section.

3.2.1.3 Clitics in Modern Irish

The particles that could be identified as clitics in Irish can be separated into a few categories:

1) **The demonstrative clitic** – Doyle (2001) notes the demonstrative particles *seo* (proximate, ‘this’), *sin* (distant, ‘that’) and *siúd* (even more distant or not even present, similar to English ‘yonder’).

a) *an bhean seo*
the woman this
‘this woman’ (Doyle 2001: 40)

b) *an fear sin*
the man that
‘that man’ (Doyle 2001: 40)

c) *an cailín siúd*
the girl yonder
‘yonder girl’ or ‘the girl mentioned’ (Doyle 2001: 40)

These morphemes appear as enclitics attaching themselves phonologically to a pronoun, noun, or noun phrase. The indiscriminacy of attachment and phonological dependence make these perfect candidates to be considered as adhering to the clitic comparative concept. However, they can also appear as independent words if they are stressed:

d) *Sin a dúirt mé.*
that say-PST I
‘That’s what I said.’ (Stenson 2020: 257)

2) **The pronominal clitic.** In Irish, pronouns have the possibility to attach (either as proclitics(2a) or enclitics(2b)) to demonstratives and nouns, like so:

a) *Is bialann maith é sin.*
BE-3SG restaurant good he that
‘That is a good restaurant.’ (Stenson 2020: 53)

b) *Is mac é.*
BE.3SG son he

‘He is a son.’ (Ó Baoill 2009: 187)

While Stenson (2020) does not specifically name these pronominals as clitics, Ó Baoill (2009) does so, and their prosodic deficiency couple with indiscriminate attachment does appear to justify Ó Baoill’s assignment. Pronominal clitics can also be emphatic, used to place special emphasis on another noun or pronoun. Doyle (2001) considers these as suffix-like enclitics *-sa/-se* (1st/2nd person singular and 3rd person singular feminine; 2nd person plural), *-na/-ne* (1st person plural), *-san/-sean* (3rd person masculine):

- c) *mo leabharsa*
my book=EMP
‘**my** book’ (Doyle 2001: 38)
- d) *do mhathairse*
your mother=EMP
‘**your** mother’ (Doyle 2001: 38)

One more particle that is considered a pronominal clitic in Irish by Ó Baoill (2009) is the enclitic element *féin*, roughly corresponding to the English *self*:

- e) *sinn féin*
our self
‘ourselves’ (Ó Baoill 2009: 187)
- f) *Déan tú féin é!*
do-IMP you self it
‘Do it yourself!’ (Stenson 2020: 159)

- 3) **The possessive clitic.** Denoting the possessive form of the noun, the particles *mo* (1st person singular), *ar* (1st person plural), *do* (2nd person singular), *bhur* (2nd person plural), and *a* (3rd person) typically appear as proclitics to the noun:

- a) *a bhéile*
POS.3.SG.M meal

‘his meal’ (Stenson 2020: 162)

b) *ár* *gcarr*

POS.1.PL car

‘our car’ (Stenson 2020: 162)

The possessives *mo* and *do* are reduced before vowels:

c) *m’athair*

my=father

‘my father’ (Stenson 2020: 162)

d) *d’iníon*

your=daughter

‘your daughter’ (Stenson 2020: 162)

There is some ambiguity here – could only the weak forms perhaps be said to be clitics? Their attachment seems to be confined to nouns only, so it is not quite clear if they can be assigned the label.

4) **Mood clitics.** To express mood Irish uses interrogative particles *an/ar* and negative particles *ní/nach* which can show clitic-like qualities like attachment to either verbs or pronouns:

a) *Nílim* *sásta.*

NEG.BE.1SG happy

‘I am not happy.’ (Stenson 2020: 88)

b) *Deir* *sé nach* *dtuigeann* *sé.*

say.3SG he NEG understand-3SG he

‘He says that he doesn’t understand.’ (Stenson 2020: 46)

c) *An* *dtuigeann* *tú?*

INT understand.2SG you

‘Do you understand?’ (Stenson 2020: 46)

The negative particle *nach* may also be used as a negative interrogative particle in place of *an/ar*:

- d) *Nach labhraíonn sí Gaeilge?*
NEG speak-3SG she Irish
'Doesn't she speak Irish?' (Stenson 2020: 44)

These particles are not noted as clitics by any scholars in the relevant literature, however, due to their clitic-like qualities they will be considered as candidates in this paper.

3.2.2 Scottish Gaelic

3.2.2.1 A historicity and overview of Scottish Gaelic

It is generally believed that Scottish Gaelic (henceforth to be referred to as 'Gaelic') is a direct descendant of Old Irish brought to Scotland by settlers from Ireland in the 4th-5th centuries, ousting the then widely used Pictish (Gillies 2009). As such, Irish (particularly Ulster Irish) and Gaelic are closely related and share many grammatical features. As time went on the use of Gaelic gradually declined, giving way to the Middle English-derived Scots, particularly in the Lowlands. Speakers of the language migrated towards the Highlands and Islands areas of Scotland, and it is in these areas where the majority of Gaelic is used today. Unfortunately, the development of the language does not have the level of attestation that Irish does, as such, it is difficult to compare the grammar of Gaelic today to the way it was used in the centuries before (Gillies 2009). However, it is likely safe to assume that due to the similar development of Irish and Gaelic their clitic-candidate particles probably do not diverge.

As an actual community language used in daily life, Gaelic is found in the Hebrides and on parts of the western seaboard. The number of fluent speakers stands at around 57,000, coming in at just around 1% of the country's population, although it is unclear just how many of these are native speakers. The language certainly does not enjoy the same kind of status in its country as Irish does, however, attempts to protect and revive it are underway.

3.2.2.2 Clitics in Scottish Gaelic

As mentioned above, due to the similarity of Gaelic grammatical features to Irish ones, most of the particles that could be identified as clitics match across the languages, with some subtle differences:

- 1) **The demonstrative clitic.** According to Gillies (2009), just like Irish, Gaelic has the demonstrative elements *seo* (proximate), *sin* (distant), and *siud* (even more distant):

a) *an fear seo*
the man this
'this man' (Gillies 2009: 267)

However, unlike Irish, these elements may fully encliticize to the noun, for example, *seo* becoming *-sa*:

b) *an fear-sa*
the man=this
'this man' (Gillies 2009: 267)

- 2) **The pronominal clitic.** Gaelic pronouns have the possibility to encliticize to a noun:

a) *Is mac e.*
BE-3SG son he
'He is a son.' (Gillies 2009)

Emphatic pronominal particles *-sa/-san/-se/-ne* exist in Gaelic as well (note that they look identical to the demonstrative clitic, although bear different meaning):

b) *mo mhac-sa*
my son=EMP
'my son' (Gillies 2009: 268)

There is also the reflexive particle *fhéin*, which corresponds to the English 'self, own':

- c) *mo thaigh fhéin*
 my house own
 ‘my own house’ (Gillies 2009: 278)

Gillies (2009) names these particles as clitics, and this view is shared by Brennan (2009) as well as Green (2001).

- 3) **The possessive clitic.** Gaelic has similar particles to Irish which denote the possessive form of the noun, *mo* (1st person singular), *ar* (1st person plural), *do* (2nd person singular), *ur* (2nd person plural), and *a* (3rd person), typically appearing as proclitics:

- a) *mo chù*
 my dog
 ‘my dog’ (Gillies 2009: 281)

Possessives *mo* and *do* are reduced before certain vowels, just like Irish:

- b) *m' Eirinn*
 my Ireland
 ‘my Ireland’ (Gillies 2009)

With this category the same ambiguity exists as with Irish – although named by Gillies (2009) as clitics, these particles do not exhibit indiscriminate attachment, as such, their clitic status is dubious.

- 4) **Mood clitics.** Gaelic shares this proposed category with Irish as well, using the interrogative particle *an*, negative particle *cha(n)*, and negative interrogative particle *nach* to convey mood:

- a) *An cuala tu mu dheidhinn Johanna?*
 INT hear.PST.DEF you about Johanna
 ‘Did you hear about Johanna?’ (Bartlett 2021: 259)
- b) *Nach fhaca tu i?*

INT.NEG SEE.PST.DEP you her

‘Did you not see her?’ (Bartlett 2021: 259)

c) *Chan eil mi toilichte.*

NEG BE-1SG I happy

‘I am not happy.’(Bartlett 2021: 268)

Mood clitics were proposed by Bartlett (2021), and it does seem like this category holds weight – as it will be seen in the Goidelic languages’ summary.

3.2.3 Manx

3.2.3.1 A historicity and overview of Manx

Manx is a language which, much like Scottish Gaelic, descended from Old Irish brought on by Irish settlers in the 4th-5th centuries. Most attestations of the older Manx variants are translated Bibles, and what they show is that Manx really split off into its own form from Middle Irish and retains most of its conservative qualities today (Broderick 2009). Clitic particles found in earlier versions of Manx are similar to those found in its modern version, although with less frequency.

Spoken exclusively in the Isle of Man, Manx is a revived language, having been declared extinct in 1974 with the death of the last speaker to grow up in a Manx-speaking community environment (Broderick 2009). Since then, great revitalization efforts have been made, and in 2010 the UNESCO Atlas of the World’s Languages in Danger began classifying it as Critically Endangered, i.e. one step above Extinct, effectively bringing it back from the dead. As per the 2021 Isle of Man census, around 2,000 people identified themselves as being fluent in Manx, constituting 2.5% of the island’s population, however, only 23 of those identified it as their first language. Although closely related to both Irish and Scottish Gaelic, Manx displays a slightly higher degree of separation from them than Irish and Scottish Gaelic do to each other due to its preservation of some of the more archaic features of Old Irish (Broderick 2009). This is likely due to the fact that existing in an island, Manx has enjoyed less direct language contact with neighboring languages than its Goidelic counterparts, allowing it to evolve differently.

3.2.3.2 Clitics in Manx

Due to a much smaller selection of Manx sources than its Goidelic counterparts (one descriptive grammar and two available articles), the search for particles that could be identified as clitics has proved more difficult. Nevertheless, it does have such examples – some of which are akin to Irish and Scottish Gaelic, and some unique.

- 1) **The demonstrative clitic.** In line with other Goidelic languages, Manx also exhibits encliticized demonstrative particles, such as *sho* (proximal):

- a) *ny* *deiney sho*
the men this
‘these men’ (Fife 2009: 17)

Since these particles exhibit the same features as their Irish and Gaelic counterparts, applying the same principle they can be considered as falling into the clitic comparative concept.

- 2) **The pronominal clitic.** Manx also has a correspondent to the English self, ‘hene’:

- a) *mee-hene*
me=self
‘myself’ (Broderick 2009: 342)

Once again, this is also very correspondent to the category of pronominal clitics in Irish and Gaelic.

- 3) **Mood clitics.** Manx exhibits a similar case of mood indicating particles that could be identified as clitics, for example the negative particle *cha*:

- a) *cha* *nel*
NEG be
‘am not/is not/are not’ (Lewin 2016)

It looks like this category is also the same as its Goidelic counterparts, however, it is difficult to make such claims due to a very low number of sources and examples found in them.

4) **The possessive clitic.** Possessiveness in Manx is expressed through particles such as *my*:

- a) *my* *hie*
my house
'my house' (Broderick 2009: 342)

5) **The tense-aspect-mood clitic.** In a somewhat unique case, Lewin (2016) claims that Manx has the tense-aspect-mood marker particle *dy/y* which is used to form a variety of sentences. It can convey the progressive tense (a), reason/purpose (b), form an adverb (c), or express the infinitive form (d) as the English 'to' in 'to do':

- a) *dy* *my chlashtyn*
PP me hear
'hearing me' (Broderick 1984: 31)
- b) *dy* *eiyr t mee*
RSN chase me
'(in order to) chase me' (Broderick 1984: 82)
- c) *dy* *mie*
ADV good
'well' (Broderick 1984: 291)
- d) *oo* *y* *eiyr t*
you to chase
'to chase you' (Broderick 1984: 82)

While Broderick himself does not label these particles as clitics, Lewin (2016) prefers to consider them so. On one hand, it is difficult to disagree, as the high degree of indiscriminate attachment and prosodic deficiency would certainly seem to imply that these could be described as clitics, at least according to Haspelmath's criteria. However, the high variability of ways it can be used does raise some questions that will be discussed in the Goidelic languages' summary.

3.2.4 A summary of results in Goidelic languages

As it has been seen, the Goidelic languages exhibit more or less the same categories of particles that could be considered clitics. Some of those categories have been proposed by scholars who comprised the respective grammars or have written an article detailing some morphosyntactic aspects of one of the languages. Other categories are proposed by this paper on the grounds that they meet or more of Haspelmath's criteria. Upon closer inspection, however, it might be found that not all of these proposed categories qualify for genuine cliticness.

The demonstrative clitics across all three languages are pretty straightforward. They fulfill the indiscriminate attachment criterion because it has been seen that they can attach to either a pronoun, noun, or noun phrase (an example of the latter in Irish - *mac maith seo/a good son, **this one***). Both phonological dependency and 'rootlessness' are there as well, as such, this case seems to not be controversial at all.

The pronominal clitics also seem to exhibit the same qualities in all three languages, although there appears to be some disagreement on whether the possessive emphatic pronominal particles (Irish 2c, 2d; Gaelic 2b) are suffixes or clitics. Indeed, not a straightforward case. Upon first glance, it appears that the emphatic particles *-sa/-se/-san/-sean* in Irish and Gaelic (and similar in Manx) are contractions or forms of the particle *seo*, in which case they would probably match the clitic comparative concept, as they likely could emphasize not only a noun, but a noun phrase as well. However, there is an exception – for the first person plural possessive *ar* the emphatic particles would be *-na/ne*, whose origin is unclear. It is possible they are a contraction of *anseo*, meaning 'here', but there is no evidence to substantiate this claim, and seemingly no reason for it being the only particle not contracting some form of *seo*. A look into a grammar of Old Irish (McCone 2005) provides no answers, as the emphatic particles there are virtually the same, with no designated origin. So, if these particles are indeed unconnected to *seo*, does it mean that they are definitely suffixes? It would seem that there is no evidence to suggest otherwise. They can appear only after the noun they emphasize, and there were no available examples of them being able to attach to anything else. If one does subscribe to the *free word—clitic—affix* continuum theory, these emphatic possessive particles would certainly gravitate more toward affixhood, as do not appear to be function words.

The proposed category of possessive clitics raises some questions as well. From the examples found in various grammars of all three languages it seems that these particles can only attach to nouns –

no other variants were found to be permissible (unlike, for example, English ‘my’, which can attach itself to different parts of speech – ‘my son’, ‘my good son’, ‘my aptly titled essay’). Only the formation of a phonological word with the following noun makes it a clitic, but that criterion alone is simply not good enough. It might be argued that the category of function words is better suited to describe these particles.

Moving on to Bartlett’s (2021) intriguing proposed category of mood clitics. It is certainly not unheard of for languages to manifest mood with the help of clitics (Spencer and Luís 2012). The Goidelic languages exhibit these particles, but can they be called genuine clitics? Well, for one thing, they do exhibit some indiscriminacy when it comes to attachment – they can attach to pronouns, adjectives, and verbs. Examples from Irish:

- a) *Níl suim agam.*
 NEG interested on-me
 ‘I am not interested.’ (Stenson 2020: 258)
- b) *Ar mhaith leat tae?*
 INT like on you tea
 ‘Would you like tea?’ (Stenson 2020: 284)
- c) *Níl sé fuar.*
 NEG it cold
 ‘It isn’t cold.’ (Stenson 2020: 181)

As seen in the Irish analysis example 4a, the particle also has a possibility to form an orthographic word with the verb ‘be’, while not being the root of said word. These qualities seem to ascertain the fact that Celtic (at the very least Goidelic) languages express mood with the help of clitics – at least according to Haspelmath’s criteria. Not appearing to have the syntax of function words, these particles can be classified as special clitics.

The tense-aspect-mood of clitics of Manx are a bit similar to mood clitics of Irish and Gaelic, though it seems they exhibit a higher degree of variability. The question to ask here is whether all *dy/y* particles are clitics. The most likely answer is no – for instance, in Manx example 4d the particle *y* has only been found to attach itself to a noun, falling out of probably the most important cliticness criterion. In fact, the cliticness of other examples is dubious as well – it would seem that they can also

only attach to specific word classes. As such, Lewin's (2016) distinction of these particles as clitics seems a bit far-fetched – they might simply be tense-aspect-mood markers.

3.3 Brythonic languages

3.3.1 Welsh

3.3.1.1 An overview of Welsh

Out of all the Celtic languages, Welsh is considered by UNESCO to be the least endangered, with the 2021 census placing the number of fluent speakers at 538,300. This number makes up 17.8% of the population of Wales, and places Welsh quite high among other Celtic languages in terms of the amount of fluent speakers – however, the census does not state if these are L1 or L2 users, so it is unclear how many of those actually use Welsh in daily life (remember the case of Irish – the number of capable speakers is 1.8 million, but actual daily users are only 70,000). Welsh is from a different branch of the Insular Celtic group than the languages discussed above – Irish, Scottish Gaelic and Manx belong to the Goidelic branch, while Welsh is the biggest Brythonic language. The Insular Celtic hypothesis postulates that what was once proto-Insular Celtic split off into Goidelic and Brythonic, with the Goidelic Old Irish moving into Ireland and Old Brythonic staying in what is now south Britain (Willis 2009). As time went on, Brythonic differentiated into a northern branch, now represented by Welsh, and a southwestern branch represented by Cornish and Breton.

Welsh is comparable to Irish in terms of its sociolinguistic status within its country, and although it is not the 'first' official language in Wales, it is quite pervasive in public life – road signs, official documents and even schools are available entirely in the medium of Welsh (Jones & Williams 2009). With plans by the Welsh government to double the number of speakers by the year 2050, it does not look like the language will remain classified as endangered much longer.

3.3.1.2 A historicity of Welsh

Welsh is typically divided into these historical periods: Primitive Welsh (6th-9th century); Old Welsh (9th-12th century); Middle Welsh (12th-14th century); Early Modern Welsh (14th century); and

Modern Welsh (15th century-present). Of the older non-modern versions Middle Welsh is the one that is attested the most, done so in a large body of texts – tales, romances, legal codes, chronicles, religious texts, scientific works, poetry, etc. (Willis 2009)

Both Old and Middle Welsh are noted by Willis (2009) as well as Borsley et al. (2009) to have had particles which may be considered clitics. Attestations as far as Old Welsh show an interesting category – the object (also called dative) clitic, intervening between verbal prefixes (a) or attaching to an additional particle if no prefix host was present (b):

- a) *Deus* *dy-m-gwares*.
 God PRT+1s+save.PRES.SUBJ.3s
 ‘May God deliver me.’ (Willis 2009: 146)
- b) *Llawrydet* *am* *dwc*.
 sadness PRT+1s seize.PRES.3s
 ‘Sadness seizes me.’ (Willis 2009: 146)

As it will be seen later, it seems that these object clitics, at least in some form, survived in Modern Welsh. However, there is a category noted by Borsley et al. (2007) found in Middle Welsh that seems to have been rendered obsolete in the modern version of the language – the accusative clitic, attaching to the end of the particle preceding a finite verb:

- c) “*Ie,*” *heb* *ef,* “*mi* *a* ‘*e* *kymeraf.*”
 yes QUOT 3MS.IND 1S.IND PRT 3S.ACC take.PRES.1S
 “‘Yes,’ he said, ‘I shall take it.’” (Borsley et al. 2007: 321)

Other clitic categories have been attested in Old and Middle Welsh, and it seems that for the most part they are present in Modern Welsh as well. The following section takes a closer look at these particles.

3.3.1.3 Clitics in Modern Welsh

The clitic categories in Welsh proposed by Roberts and Shlonsky (1996), Mittendorf and Sadler (2011), King (2003), and Borsley et al. (2007) are comparable to those in other previously discussed

Celtic languages. However, they feature slightly different attachment as well as some other unique features such as echo pronouns.

- 1) **The pronominal clitic.** According to Roberts and Shlonsky (1996), much like its other Celtic counterparts, Welsh exhibits pronominal enclisis, which may appear in a few different forms, for example, the infix pronoun in (1a). These pronouns are enclitic to sentential particles:

a) *Mi'ch gwelais chwi.*
 PRT=you see.PST.1S you.
 'I saw you.' (Roberts & Shlonsky 1996: 182)

The researchers also define a rather unique feature of Welsh called echo pronouns – suffixed clitic pronouns which repeat information given earlier in the phrase:

b) *Mae Megan wedi ei weld o.*
 BE.3SG Megan after his see he
 'Megan has seen him.' (Roberts & Shlonsky 1996: 182)

c) *Dyweddod y bachgen ein bod ni wedi cyrraedd.*
 say.PST the boy our be.PRG us after arrive
 'The boy said that we had arrived.' (Roberts & Shlonsky 1996: 182)

The syntactic distribution of these particles is quite similar to the one found in Goidelic languages. The fact that they exhibit doubling presents an even stronger case for cliticness.

- 2) **The possessive/genitive clitic.** Several researchers note these particles that appear in possessive constructions.

a) *ei wraig*
 his wife
 'his wife' (King 2003: 307)

b) *fy nghar i*
 1s. car me

‘my car’ (Borsley et al 2007: 158)

Possessive proclitic pronouns must also be repeated with additional genitive case nouns, with the repetition encliticizing to the preceding conjunction. This is a case of clitic doubling:

c) *ei* *ffagots* *a’i* *phys hi*
her faggots and=3SF peas 3SF
‘her faggots and peas’ (Mittendorf & Sadler 2011: 387)

It is also possible for these particles to completely encliticize to preceding certain vowel-final words, especially prepositions:

d) *i’w* *dŷ*
to=3.SG.M house
‘to his house’ (Borsley 2022: 4)

This proposed clitic category is a bit different from its Goidelic counterpart. These particles seem to exhibit a higher degree of indiscriminate attachment as well as doubling, features unique to Welsh. This makes them more considerable for genuine clitic-hood, and will be discussed in the Brythonic languages summary.

3) **The determiner clitic.** Welsh has been observed by Borsley et al. (2007) as well as Mittendorf and Sadler (2011) to use certain determiners as clitics, for example, the definite article, which can appear both separately (a) and fully encliticized (b):

a) *y* *tair* *cath* *ddu*
the three.F cat black
‘the three black cats’ (Borsley et al. 2007: 152)

b) *y* *dynion a’r* *merched*
the men and=the girls
‘the men and the girls’ (Mittendorf & Sadler 2011: 405)

Example (b) again showcases Welsh clitic doubling – in a sentence with two or more definite NPs definite articles must be repeated.

- 4) **The object-agreement clitic.** Borsley et al. (2007) proposes this clitic category. Typically found in *wh*-questions, this particle precedes the verb. It is usually present in literary Welsh but has been observed to be omitted in colloquial Welsh. It normally agrees in person and number with the *wh*-object:

a) *Beth ydych chi'n (ei) fwyta?*
what be.PRES.2P you=PROG (3MS) eat.INF
'What are you eating?' (Borsley et al. 2007: 110)

Using object-agreement clitics in non-*wh*-environments allows clitic doubling as well (this is also optional):

b) *Mae Ifan yn ei fwyta (e).*
be.PRES.3S Ifan PROG 3MS eat.INF (it)
'Ifan is eating it.' (Borsley et al. 2007: 110)

This is an ambiguous case. The researcher describes these particles as always preceding the verb, and virtually non-existent if there is no verb for it to attach to. As it will be seen, other Brythonic languages exhibit similar particles. Comparing and discussing in the Brythonic language summary will help see where they stand.

3.3.2 Breton

3.3.2.1 An overview of Breton

Breton is the only Celtic language spoken in mainland Europe, despite being classified as Insular Celtic instead of Continental Celtic. This is because the language originated in Great Britain and was brought over to Armorica (present-day Brittany in France) by migrating Britons (Press 2009). The current number of speakers is placed at around 210,000 by the 2018 census, placing it as the

second most spoken Celtic language, however, just like with the Welsh census, it is unclear whether this number refers to only L1 speakers or if L2 speakers are included, therefore, no generalizations can be made. The UNESCO *Atlas of the World's Languages in Danger* classifies Breton as ‘severely endangered’ due to the aging majority of speakers (many being over the age of 60) as well as the fact that the number of speakers decreased by 80% in the last 60 years. Breton is the only living Celtic language that is not recognized by a national government as an official or regional language (Timm 2009). Despite these challenges, a revitalization effort is underway, with the number of children attending bilingual classes rising 33% between 2006 and 2012.

3.3.2.2 A historicity of Breton

The recognized historical stages of Breton are Old Breton (9th-12th century); Middle Breton (12th-17th century); and Modern Breton (17th century-present). It is not as widely attested as older versions of Welsh, but the smaller corpus still allows for some linguistic analysis (Press 2009).

Breton is concordant with earlier versions of Welsh in the clitic particles that are found in the languages. Just like Middle Welsh, Rezac (2020) identifies object clitics in Middle Breton, however, subject clitics can be found in some perfective constructions as well (*o* – subject proclitic; *ma* – object proclitic):

- a) *pere o=deueus ma=dibilitet*
 which 3PL=have 1SG=weakened
 ‘which have weakened me’ (Rezac 2020: 2)

The dative version of the object clitic may appear infix as well:

- b) *An=guen heguen a~m~louenas*
 the=white smiling R=1SG=gladden.PT
 ‘The smiling white one gladdened me.’ (Rezac 2020: 12)

Interestingly enough, the infix version of the particle did not survive in Modern Breton. However, some other particles described as clitics did, and they will be discussed in the following section.

3.3.2.3 Clitics in Modern Breton

The clitic categories observed by Press (2009), Jouitteau and Tamarit (2023), Stump (1984), and Weisser (2019) in Modern Breton are unremarkable in the context of current Celtic languages. As such, a deeper analysis of them will be offered in comparison in the Brythonic language summary. However, it seems that the language had quite a few more categories in its earlier iterations, most notably Middle Breton. In its attestations, scholars have noted such particles as dative, locative, genitive, and accusative clitics, none of which supposedly remain today (Rezac 2020).

- 1) **The pronominal clitic.** Breton weak pronouns have the typical Celtic feature of pronominal cliticization. All pronouns have a weak form, which manifests in proclisis and enclisis (Press 2009):

	strong	proclitic	enclitic
1PS	<i>me</i>	<i>am-em- 'm/va- ma</i>	<i>-me</i>
2PS	<i>te</i>	<i>az-ez- 'z/da</i>	<i>-te</i>
3PSf	<i>hi</i>	<i>he/hec 'h</i>	<i>-hi</i>
3PSm	<i>eñ</i>	<i>e-en</i>	<i>-eñ</i>
1PP	<i>ni</i>	<i>hon/hor/hol</i>	<i>-ni</i>
2PP	<i>c 'hwi</i>	<i>ho/hoc 'h</i>	<i>c 'hwi/-hu</i>
3PP	<i>i, int</i>	<i>o</i>	<i>-i(nt)</i>

It is noted by Weisser (2009) that weak clitic pronouns are used emphatically, much like they can be used so in Irish and Scottish Gaelic:

- a) *Levrioù a lennan-me.*
 books PRT read.1SG=EMPH.1SG
 ‘I read books.’ (Weisser 2019: 299)

In a feature similar to Welsh, Breton has also been observed by Jouitteau and Tamarit (2023) to make use of the echo pronoun, which, for example, may encliticize to an adjective:

- b) *e di bihan-eñ war ar maze*
his house small.3SGM on the country

‘his beautiful house in the countryside’ (Jouitteau & Tamarit 2023: 17)

- 2) **The possessive clitic.** Possession in Breton can be indicated by means of a proclitic possessive pronoun:

- a) *e dad*
his father

‘his father’ (Press 2009: 441)

With the possessive clitics, emphasis can also appear, with a pronominal enclitic appearing with a corresponding proclitic:

- b) *e dad-eñ*
his father=EMPH

‘**his** father’ (Stump 1984: 344)

- 3) **The determiner clitic.** Demonstratives in Breton can be conveyed by the attachment of enclitics, similarly to the ones found in Goidelic languages - *-mañ* (proximate, here), *-se* (distant, there), and *-hont* (very distant from the speaker or not in view at all, over there, yonder).

- a) *ar stêr-mañ*
the river=here

‘this river’ (Press 2009: 455)

- b) *al lent-hont*
the lake=yonder

“that lake over there” (Fife 2009: 17)

4) **Object clitics.** Similarly to Welsh, Breton has been recorded by Joutiteau and Tamarit (2023) to exhibit object clitics as well. They typically appear proclitic to verbs:

- a) *Me* ‘ *meus* *e* *ziskouezet* *em* *levr.*
I PRT 1SG.has 3SGM show.PPF in-my book
‘I have described him in my book.’ (Joutiteau & Tamarit 2023: 12)

Their use appears to be optional and is even considered to be archaic (Joutiteau and Tamarit 2023). Here, similarly to the object-agreement category in Welsh, the proposed cliticness is dubious, as these particles are only recorded to attach to verbs.

3.3.3 Cornish

3.3.3.1 An overview of Cornish

The Cornish language, spoken exclusively in the United Kingdom County of Cornwall, is similar to Manx in the sense that it is a revived language, having once gone extinct and having since been revived. However, the revival of Cornish proved to be a bit more difficult than that of Manx due to the amount of time passed since the recorded extinction – while the revivors of Manx had virtually a few decades of extinction to deal with, Cornish had fallen out of use by mid-18th century, with the last capable speaker dying in the 19th century. As such, the revival effort, which started at the beginning of the 20th century, features conflicting views from scholars as to which variety of the language should be considered for the revived version. The founder of the movement, Cornish cultural activist Henry Jenner published a handbook of the language based on Late Cornish (17th-18th century), however, a much more popular work outlining an orthography called *Unified Cornish* was published by Robert Morton Nance in 1929, drawing on Middle Cornish, spoken in the 14th and 15th centuries. The author asserted that this period represented the high point of Cornish literature, and as such, should be taken as the standard. As the 20th century went on, other scholars have had their say presenting proposed Cornish orthographies, and by the 21st century, there were as many as five different orthographies in use. It became clear that such chaos is unsustainable for the revival effort, and, as such, the *Gorsedh Kernow*, an organization devoted to maintaining the national Celtic spirit of Cornwall, adopted a

variety known as Standard Written Form in all official matters concerning Cornish in 2009. The variety in question is based on a compromise of all other varieties, its grammar and orthography being a mixture of Middle Cornish and Late Cornish (George 2009). In the same year Standard Written Form was adopted, UNESCO *Atlas of the World's Languages in Danger* changed the status of Cornish from Extinct to Critically Endangered, marking what some believe to be the official point of the language's return from the dead. According to the 2021 census, there are no speakers of Cornish to whom it would be the first language, however, 563 were marked as being fluent L2 speakers.

3.3.3.2 A historicity of Cornish

Cornish is usually divided into five phases: Primitive Cornish (7th-9th century); Old Cornish (9th-13th century); Middle Cornish (13th-16th century); Late Cornish (16th-19th century); and Revived Cornish (20th century-present) (George 2009). The last category is somewhat contentious, as it is not exactly a new development of the language but rather a reimagining of it using elements from both Middle and Late Cornish. Unsurprisingly, these are the most attested versions of the language, with Middle Cornish arguably boasting the largest body of texts. Since Cornish ultimately does offer development past the level of Late Cornish, the possible clitic variants analyzed in the next section come from the last two iterations of the language.

3.3.3.3 Clitics in Cornish

Cornish proved to be the most difficult Celtic language to analyze clitics in due to a variety of proposed standards. While Standard Written Form is the currently accepted standard of Cornish for official matters, George's (2009) description of the grammar is the most exhaustive, which is mostly a comparison of Middle Cornish and Late Cornish. As such, possible clitic particles from both these varieties will be compared.

- 1) **The pronominal/possessive clitic.** As noted by George (2009), in both Middle and Late Cornish personal pronouns have strong and weak forms, the latter of which feature encliticization. In Middle Cornish, clitic pronouns were a feature of emphasis, which also exhibited clitic doubling:

- a) *ow* *hath* *vy*
my cat I
'my cat' (George 2009: 515)

Meanwhile only the possessive clitic *ow* is used when there is no special emphasis:

- b) *ow* *hath*
my cat
'my cat' (George 2009: 515)

In Late Cornish the possessive clitic *ow* is replaced by the definite article *an*, while personal pronoun *vy*(I, me) becomes a marker for possessiveness:

- c) *an* *gath* *vy*
the cat me
'my cat' (George 2009: 516)

The last part – replacing a possessive clitic with a definite article – is unique to Cornish. Due to a small sample of analyzed examples, it is difficult to make any strong generalizations about this, but it is interesting that possessiveness is expressed with a nominative pronoun following the noun. It would probably mean that in Late Cornish, possessive clitics ceased to exist. And Standard Written Form, the variety used in all matters official today, appears to use the Late Cornish possessive version.

- 2) **The determiner clitic.** A bit similar to the other Celtic languages, Cornish also has clitics for demonstrative determiners. But instead of the examples 'here', 'there', and 'yonder' seen in previous analyzed languages, George (2009) notes that Cornish produces the demonstratives 'this/these' and 'that/those' with the help of the clitics *ma* and *na*:

- a) *an* *lyver* *ma*
the book this
'this book' (George 2009: 514)

- b) *an lyvrow na*
 the books those
 ‘those books’ (George 2009: 514)

The definite article *an* may also fully encliticize after a vowel in certain combinations:

- c) *ha'n lyvrow*
 and=the books
 ‘and the books’ (George 2009: 498)

- 3) **Object clitics.** In concordance with other Brythonic languages, particles identified as object clitics have been identified by Rezac (2020) in Cornish as well. An example from Middle Cornish below (exhibiting clitic doubling as well):

- a) *ha ty a=vyʒ=hy*
 and 2SG R.2SG=be.CNS=3SGF
 ‘and you will have her’ (Rezac 2020: 18)

However, it is unclear if these particles survived into Late Cornish, as no examples were found.

3.3.4 A summary of Brythonic languages results

The Brythonic languages exhibit some similar proposed clitic categories to their Goidelic counterparts. As it has been seen, determiner clitics in Cornish and Breton are very similar to those found in Irish, Gaelic, and Manx. Since their clitic status has already been discussed in the summary of Goidelic languages, that will not be necessary here.

However, with the case of determiner clitics, the Welsh inclusion of the definite article merits discussion. Borsley et al. (2007) claims that the definite article is a clitic because numerals and other elements may intervene between it and the noun:

- a) *y tair cath*
 the three cats

‘the three cats’ (Borsley et al. 2007: 151)

- b) *Yr un bachgen ydyw.*
the same boy BE-3SG

‘He is the same boy’ (Borsley et al. 2007: 155)

But as it was seen in Welsh example 3b, there is a need for the definite article to be repeated before each definite noun. Can one definite article attach to a whole definite noun phrase with multiple nouns? No such examples were found. And such was the case across all Celtic languages. Would it be fair then to consider the Welsh definite article a clitic? Another argument for it might be that it may become enclitic to a preceding word and takes the form *r* if that word ends in a vowel:

- c) *o’r tŷ*
from=the house

‘from the house’ (Borsley et al. 2007: 155)

In this example the definite article attaches to a preceding preposition. Keeping in mind Haspelmath’s criteria, that certainly would mean that the Brythonic definite articles gravitate towards cliticness more so than the Goidelic ones might as they offer more attachment possibilities.

The pronominal clitics present a far less controversial case – as they did in the Goidelic languages analysis. Weak pronoun forms may attach to a variety of parts of speech, are phonologically reduced, and may be subject to doubling – all criteria are fulfilled.

Possessive particles in Brythonic languages appear to be a bit different than those in their Goidelic counterparts. As seen in Welsh examples 2b and 2c, the particles may attach to conjunctions and prepositions instead of the nouns whose possessiveness they mark (although Breton and Cornish examples offer little conclusiveness). This implies a higher degree of indiscriminate attachment than the Irish, Gaelic, or Manx possessive particles. Perhaps then these particles are closer to cliticness – certainly if we adhere to Haspelmath’s criteria we can describe them as matching the clitic comparative concept.

When it comes to object clitics, all three Brythonic languages were recorded to exhibit their attachment to verbs only. Even though this category was proposed by three different sources (Borsley et al. 2007; Rezac 2020; Jouitteau & Tamarit 2023), the evidence for genuine cliticness is simply not there, at least not according to the criteria set by this paper.

4. Conclusion

Having described the possible clitic particles in both Goidelic and Brythonic languages, it is possible to draw some conclusions about clitics in Celtic languages in general.

Firstly, it must be remembered that clitics are not a pre-determined category. With the number of different takes from scholars on the definition of clitic, it would be unwise to deem it so. Clitics must be understood as a comparative concept – something defined for the specific purpose of language comparison, and not a natural category of language. One might use it when attempting to ascribe properties to a particle which may not fit into any other category. But a clitic in one language will not necessarily be a clitic in another when compared or translated. This is where language-specific descriptions come in and are used by typologists.

Regarding possible clitic particles in Celtic languages, it is clear that they share some very similar categories. The pronominal clitic was attested by most scholars in all languages, to no surprise – Indo-European languages typically possess this category. The other categories were similar as well, however some definitely gravitated more toward cliticness than others. The Brythonic definite articles appeared to attach to more parts of speech than the Goidelic ones. Also, the possessive clitics in Brythonic languages exhibited a less constrained syntactic placement than the ones in Goidelic languages. One conclusion might be drawn that Brythonic languages exhibit a syntax that flows a bit more freely, which is more concordant to surrounding Germanic and Romance languages. This could be due to a higher level of language contact with neighboring languages than that experienced by Goidelic languages – geographically they are a bit more isolated than their Brythonic counterparts.

This is by no means a definitive list of particles that can be considered as clitics in Celtic languages. Other scholars might identify other language-particular descriptions of clitics that match the underlying comparative concept. But the ones analyzed in this paper are certainly those that have been identified so far, with some new suggestions as well, such as the mood clitic category in Irish and Manx.

Regarding definitions themselves, we have seen that opinions of scholars may diverge. Pronominal particles in Irish are defined as clitics by Ó Baoill (2009), but not so by Stenson (2020). Definite articles have been observed to be clitics in Welsh by Mittendorf and Sadler (2011), Borsley et al. (2007), and Awbery (2009), but are not defined so by King (2003), who in turn mentions other clitics such as the possessive one, which by itself is not recognized as such by other scholars. Are any of them wrong? Perhaps all of them? Or none? In this paper we have applied Haspelmath's criteria of

clitic identification, and ascribed particles to be either close enough to his proposed comparative concept of clitics to be determined as such, or not. But the beauty of linguistic typology is that comparative concepts are simply methodological tools, not actual linguistic categories. A comparative concept can be applied to an individual language to determine some properties of a linguistic structure, but it is not used to make definitive statements. So, by applying the comparative concept of clitics to particles in Celtic languages, we define some properties of these particles as particular to clitics, and we may even consider them as such, but that certainly does not invalidate their definitions proposed by other scholars. It simply means that there are different ways of looking at the same linguistic structures. This study defined some typical structures identified as clitics in Celtic languages, and the results might be used for a better understanding on how these particles work in a broader Indo-European context, perhaps by comparing their functionality with proposed clitics in other languages.

One of the last chapters in Spencer and Luís' book *Clitics: An Introduction* (2012) is aptly rhetorically titled *Do clitics exist?* Indeed, this is a question that may arise when delving deeper into attempting to identify this elusive phenomenon. In the chapter the authors reach the conclusion that “while the CATEGORY of clitic may not exist, some sort of CONCEPT of clitic remains ubiquitous, both in theoretical discussion and in descriptive studies, as an umbrella term” (2012: 327). This agrees with Haspelmath's comparative concept theory as well and remains a recommended conclusion to keep in mind when considering what is or is not clitic.

The lack of a clearly defined clitic concept certainly leaves a lot of room for further typological research. Regarding the results of this paper, they can be used to further research on Europeanisms and SAE – Standard Average European, a grouping of European languages with shared common linguistic features proposed by American linguist Benjamin Whorf. Regarding clitic research in general, there are still many open questions that could be addressed by further research, such as a deeper investigation into the proposed continuum between free words and clitics, or a study on how the notion of clitics affects the way morphosyntactic features are compared across languages. It seems that the field of linguistic typology has ways to go before a full and clear understanding of cliticness is achieved.

5. Data sources

- Awbery, Gwenllian. 2009. Welsh. In *The Celtic Languages* (2nd ed., pp. 359–426). London, New York: Routledge.
- Bartlett, Tom. 2021. Interpersonal grammar in Scottish Gaelic. In: Martin, J.R., Quiroz, B. and Figueredo, G. (eds.) *Interpersonal Grammar: Systemic Functional Linguistic Theory and Description* (pp. 257-284) Cambridge: Cambridge University Press. ISBN 9781108493796 (doi: 10.1017/9781108663120.009)
- Borsley, Robert D., Tallerman, Maggie, & Willis, David. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Broderick, George. 1984. *A handbook of Late Spoken Manx* (Vol. 1). Tübingen: Niemeyer.
- Broderick, George. 2009. Manx. In *The Celtic Languages* (2nd ed., pp. 305–356). London, New York: Routledge.
- Doyle, Aidan. 2001. *Irish*. Munich: LINCOM EUROPA.
- Fife, James. 2009. Typological aspects of the Celtic languages. In *The Celtic Languages* (2nd ed., pp. 4–21). London, New York: Routledge.
- George, Ken. 2009. Cornish. In *The Celtic Languages* (2nd ed., pp. 488–536). London, New York: Routledge.
- Gillies, William. 2009. Scottish Gaelic. In *The Celtic Languages* (2nd ed., pp. 230–304). London, New York: Routledge.
- Jouitteau, Melanie, & Tamarit, Francesc J. P. 2023. The Syntax of Modern Breton. *Handbook of Celtic Linguistics and Languages*. Palgrave.
- King, Gareth. 2003. *Modern Welsh: a comprehensive grammar*. London, New York: Routledge.
- Lewin, Christopher. 2016. The syntax of the verbal noun in Manx Gaelic. *Journal of Celtic Linguistics*, 17, 147–238.

- McCone, Kim. 2005. *A first Old Irish grammar and reader*. Naas: The Leinster Leader.
- Mittendorf, Ingo, & Sadler, Louisa. 2011. Welsh Prenominals: at the Syntax-Morphology Interface. *Empirical Issues in Syntax and Semantics*, 8, 383–408.
- Press, Ian. 2009. Breton. In *The Celtic Languages* (2nd ed., pp. 427–487). London, New York: Routledge.
- Rezac, Milan. 2020. Mihi est from Brythonic to Breton I. *Indogermanische Forschungen*, 125(1), 313–362. doi:10.1515/if-2020-013
- Roberts, Ian, & Shlonsky, Ur. 1996. Pronominal enclisis in VSO languages. *The Syntax of the Celtic Languages*, 171–199. doi:10.1017/cbo9780511586279.006
- Stenson, Nancy. 2020. *Modern Irish: A Comprehensive Grammar*. London, New York: Routledge.
- Stifter, David. 2009. Early Irish. In *The Celtic Languages* (2nd ed., pp. 55–116). London, New York: Routledge.
- Stump, Gregory. T. 1984. Agreement vs. incorporation in Breton. *Natural Language and Linguistic Theory*, 2(3), 289–348. doi:10.1007/bf00133790
- Timm, Lenora A. 2009. Language, culture and identity in Brittany. In *The Celtic Languages* (2nd ed., pp. 712–752). London, New York: Routledge.
- Willis, David. 2009. Old and Middle Welsh. In *The Celtic Languages* (2nd ed., pp. 117–160). London, New York: Routledge.
- Weisser, Philipp. 2019. A Remove-based theory of the complementarity effect in Breton. *Linguistische Arbeitsberichte* 94
- Ó Baoill, Donall P. 2009. Irish. In *The Celtic Languages* (2nd ed., pp. 163–229). London, New York: Routledge.

6. References

- Anderson, Stephen R. 2005. *Aspects of the Theory of Clitics*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.
- Bond, Oliver. 2018. Canonical Typology. *The Oxford Handbook of Morphological Theory*, 408–431. doi:10.1093/oxfordhb/9780199668984.013.26
- Borsley, Robert D. & Roberts, Ian. eds. 1996. *The syntax of the Celtic languages: a comparative perspective*. Cambridge: Cambridge University Press
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press
- Daniel, Michael. 2010. “Linguistic typology and the study of language.” *Oxford Handbooks Online*, <https://doi.org/10.1093/oxfordhb/9780199281251.013.0004>.
- Dixon, R. M. W., & Aikhenvald, Alexandra. Y., eds. 2002. *Word: a cross-linguistic typology*. Cambridge: Cambridge University Press.
- Franks, Steven L. 2000. Clitics at the interface. *Linguistik Aktuell/Linguistics Today*. doi:10.1075/la.30.03fra
- Graffi, Giorgio. 2010. “The pioneers of linguistic typology: From Gabelentz to Greenberg.” *Oxford Handbooks Online*, <https://doi.org/10.1093/oxfordhb/9780199281251.013.0003>.
- Halpern, Aaron. 1998. Clitics. In Spencer, Andrew, and Zwicky, Arnold (eds.), *The Handbook of Morphology*. Oxford: Blackwell Publishers.
- Haspelmath, Martin. 2007. Pre-established categories don’t exist: Consequences for language description and typology. *Linguistic Typology*, 11(1). doi:10.1515/lingty.2007.011
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in Crosslinguistic Studies. *Language*, 86(3), 663–687. doi:10.1353/lan.2010.0021

- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica*, 45(1). doi:10.1515/flin.2011.002
- Haspelmath, Martin. 2015. Defining vs. diagnosing linguistic categories: A case study of clitic phenomena. *How Categorical Are Categories?*, 273–304. doi:10.1515/9781614514510-009
- Haspelmath, Martin. 2023. Types of clitics in the world’s languages. *Linguistic Typology at the Crossroads*, 3(2), 1–59. <https://doi.org/10.6092/issn.2785-0943/16057>
- Jones, Robert O., & Williams, Colin H. 2009. The sociolinguistic context of Welsh. In *The Celtic Languages* (2nd ed., pp. 650–711). London, New York: Routledge.
- Miller, Christine. 2020. *Human Biology: Human Anatomy and Physiology*. Pressbooks, *Pressbooks*, humanbiology.pressbooks.tru.ca/, Accessed 14 Feb. 2024.
- Ó hIfearnáin, Tadgh. 2015. Irish-speaking society and the State. In *The Celtic Languages* (2nd ed., pp. 539–588). London, New York: Routledge.
- Ramat, Paolo. 1987. *Linguistic Typology*. Berlin, New York: Walter de Gruyter
- Ramat, Paolo. 2010. ‘The (Early) History of Linguistic Typology’, *Oxford Handbooks Online*. doi:10.1093/oxfordhb/9780199281251.013.0002.
- Song, Jae J. 2010. ‘Setting The Stage’, *Oxford Handbooks Online*. doi:10.1093/oxfordhb/9780199281251.013.0001.
- Spencer, Andrew & Luis, Anna R. 2012. *Clitics: An introduction*. Cambridge: Cambridge University Press.
- Velupillai, Viveka. 2012. *An Introduction to Linguistic Typology*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Zwicky, Arnold M. 1977. *On clitics*. Bloomington: Indiana University Linguistics Club
- Zwicky, Arnold M., and Pullum, Geoffrey K. 1983. Cliticization vs. inflection: English *n’t*. *Language*, 59, 502–13.

7. Summary in Lithuanian

Šiame magistro moksliniame darbe nagrinėjamas lingvistinis reiškinyje vadinamas "klitikais". Pagal paprasčiausią ir labiausiai paplitusį apibrėžimą klitikai yra trumpos kalbinės formos, turinčios sintaksinių žodžio savybių, bet fonologiškai priklausančios nuo kito žodžio ar frazės. Tačiau mokslininkai dažnai nesutaria dėl tikslios šio reiškinio apibrėžties, o kartais net dėl jo egzistavimo. Šiame straipsnyje trumpai apžvelgiama apie šį terminą atliktų tyrimų istorija (Haspelmath 2015, 2023; Spencer ir Luis 2012; Zwicky 1977; Zwicky ir Pullum 1983), taip pat bandoma viską susiaurinti iki vienos apibrėžties, taikytinos tipologinėje analizėje. Nustatytas apibrėžimas bus naudojamas klitikams keltų kalbų šeimoje, t. y. airių, škotų gėlų, valų, manksų, bretonų ir kornų kalbose, nustatyti. Ankstesni tyrimai parodė, kad šiose kalbose minėtas reiškinys pasireiškia skirtingu mastu, tačiau iki šiol nebuvo atlikta išsami šio reiškinio tarpkalbinė analizė keltų kalbose. Tuomet rezultatai bus analizuojami ir sudaryta galimų keltų kalbų klitikų dalelyčių tipologija. Bus daromos išvados dėl šios kalbinės formos vartosenos, taip pat išvados dėl tolesnių lingvistinės tipologijos tyrimų, susijusių su šia neaiškia ir kartais prieštaringai vertinama klitikų tema.