Gledhill, Christopher 2014c. Phraseology as a Measure of Emergent Norm: the Case of Esperanto. In José-Carlos Herreras (ed.), *Politiques linguistiques et langues autochtones d'enseignement dans l'Europe des vingt-sept.* Presses Universitaires de Valenciennes, pp. 317-348.

PHRASEOLOGY AS A MEASURE OF EMERGENT NORM: THE CASE OF ESPERANTO

Christopher GLEDHILL
Université Paris Diderot – Paris 7
EA 3967 - CLILLAC-ARP

1. Introduction

In this paper I discuss the problems involved in prescribing and describing the language norms of Esperanto, one of the best-known examples of a 'constructed' or 'international auxiliary' language (IAL). Although there have been many discussions about the standardisation of Esperanto in general linguistics (Blanke 1985, Savatovsky 1989, Schubert 1989, Fiedler 2006, Burkina 2009) and among Esperantists (Zamenhof 1936, Duc Goninaz 1984, Piron 1986, 1989, Wells 1989), in this paper I focus primarily on how the norms of the language manifest themselves in one particular area of language: phraseology. In the linguistics literature, there have been several studies on the grammar of Esperanto (Sherwood 1982, Duc Goninaz 1984, Daŝgupta 1989, Li 1996, Dankova 1997, Gledhill 2000, Jansen 2012, 2013) as well as phonology and pronunciation (Versteegh 1993, van Oostendorp 1999, Bergen 2001, Lindstedt 2006, Burkina 2009). However, there have to to date only been a handful of studies which focus on phraseology (Daŝgupta 1993, Fiedler 2007) and even fewer from the point of view of corpus-based linguistic analysis (Gledhill 2008, 2010).

The view of phraseology I adopt in this paper is different from the traditional approach, which concentrates on proverbs, idioms and the like. The perspective adopted here is that of descriptive (corpus-based) linguists such as Sinclair (1991) and functional linguists such as Halliday (1985) and Tucker (2007). According to this perspective, no phenomenon of language can be effectively analysed out of context, using introspection or logic. Rather, these linguists maintain that all language is 'language in use', and therefore bound to a particular 'context of situation' in the form of a 'text' (a written or spoken communication). It follows from this approach that language phenomena can only be effectively analysed through the observation of a representative corpus of texts. Furthermore, the phenomenon of **phraseology** does not just concern a certain subset of idiosyncratic expressions, but includes any meaningful combination of words, as long as it is recognised that all words enter into more or less predictable patterns of co-selection ('collocations'), and that all texts are made up of more or less pre-constructed sequences of words ('lexicogrammatical patterns'). In this paper, I argue that it is only by observing the phraseology of a

language in this way, that it is possible to observe its 'implicit' norms.

But before looking at the phraseology of Esperanto, it is important to understand where Esperanto stands in the overall typology of constructed languages (also known as an 'artificial' or 'invented' languages). A broad distinction is usually made between 'philosophical language' and 'planned language' schemes (Blanke 1985, Schubert 2001).[1] The authors of **philosophical languages** typically attempt to devise a rational system of meaning using language-neutral symbols (icons, mathematical signs, musical notes, etc.). Such schemes tend to assume that syntax is either intuitive, or works according to the rules of logic or some other formal system. Examples include John Wilkins' 1668 *Real Character*, François Sudre's 1827 *Solresol*, and James C. Brown's 1955 *Loglan*. The creators of **planned languages**, on the other hand, propose a rationalised version of an existing natural language, or a hybrid form of several languages. These schemes usually prescribe a regular system of phonology, lexis and grammar, following similar principles of simplification as pidgin or creole languages. There is also sometimes an attempt to rationalise the system of meaning, although this often amounts to avoiding polysemy or homophony in a vocabulary list. A further distinction can be made between 'national' planned languages (Dante's *De Vulgari eloquentia,* Modern Hebrew and New Norwegian) and 'neutral' planned languages (Johan M. Schleyer's 1879 *Volapük*, Ludwik Zamenhof's 1887 *Esperanto* or Alexander Gode et al.'s 1951 *Interlingua*). It is no accident that both types became prevalent during the late 19th and early 20th centuries. At this time, various peoples felt the need to adopt a national language or to standardise a previously non-standard vernacular. During the same period, the early proponents of a neutral language, or an 'International Auxiliary Language' (IAL), argued that competition between the major national (and also at the time, imperial) languages constituted a barrier to peaceful co-existence and scientific communication. Ironically, the spread of *English as a lingua franca* (ELF) has not reduced the need for, or interest in, planned languages. On the contrary, the spread of ELF has engendered a complex set of new languages, including several regional varieties English (*Singlish,* etc.), as well as controlled languages such as *Basic English, Simplified Technical English,* and so on.

Virtually alone among constructed languages, Esperanto is the only planned IAL that has developed a sizeable, autonomous speech community. This development is surprising given that, unlike most other minority languages or vernacular forms of speech, the language first came to the world in the form of blueprint (a set of principles rather than a fully planned language), and has a speech community that is largely non-native, self-taught and unregulated. In the first sections of this paper, I present an overview of the problems of planning the norms of such a language, as well as some of the broader issues raised by the study of IALs. In particular, I set out how the norm of Esperanto has evolved from the 'explicit' statements of its inventor and early followers into an 'implicit' norm, that is to say a state of language use which enjoys a high degree of internal coherence. As I argue in the final sections, this coherence manifests itself in the typical phraseology of the language, as can be

[1] Various other types exist, including computer programming languages, fictional and fantasy languages, secret codes and other pseudo-languages (c.f. Yaguello 2006, Okrent 2009, Albani & Buonarroti 2010, Romaine 2011).

observed using the methods of computer-based corpus linguistics.

2. The Esperanto movement as a speech community

There now exists an extensive body of sociolinguistic work on Esperanto as a social phenomenon, and Esperantism as a political and ideological movement (Forster 1982, Large 1985, Comrie 1996, Stocker 1996, Smith 2011). Most observers agree that the status of Esperanto has evolved over time: in 1887, the date of Ludwik Zamenhof's first attempt to publicise the language, Esperanto consisted of little more than a 40-page pamphlet. Yet within a few decades, Esperanto had acquired much of the infrastructure of a living natural language. The following list presents a summary of the main features of the contemporary Esperanto movement:

1. Speech community:
- Estimates vary between 500,000 and 3,500,000 **second-language (L2) users** of varying degrees of competence (Pool & Grofman 1989, Edwards 1994, Fiedler 2006). A more reliable figure for committed L2 users comes from the *Universala Esperanto Asocio,* which currently has 15,000 members (Buller 2013),
- Approximately 1,000 'mother-tongue' **first-language (L1) users** (as reported in Versteegh 1993, Corsetti 1996, Bergen 2001, Lindstedt 2006).

2. Acceptance (official sources of a symbolic, institutional language norm):
- Some **official recognition** by organisations such as the International Telegraphic Union (1938), UNESCO (1954), the Vatican (1990) and others (c.f. Korĵenkov 2005),
- A variety of national and international Esperantist **organisations** using Esperanto as their working language, the best known being the *Universala Esperanto Asocio* ('Universal Esperanto Association' or UEA, based in permanent premises in Rotterdam), but also other more specialised organisations (*Sennacia Asocio Tutmonda* 'Worldwide Anti-nationalist Association', *Kristana Esperantista Ligo Internacia,* 'International League of Christian Esperantists', etc.)
- Many **local meetings** in many major cities throughout the world. These often take place in premises owned or rented by the national association,
- A modest presence in some **national education systems**, supported by the existence of exams and certification organised by Esperantist associations (Kováts 2007).

3. Codification (official sources of a deliberate, prescriptive language norm):
- A **standard** for the language proposed by Zamenhof (1887) and then ratified in the form of the *Fundamento de Esperanto* 'Foundation of Esperanto' (Zamenhof 1905),
- A **language academy** *Akademio de Esperanto* concerned with defending the original scheme and making judgments on as yet unprescribed (cf. Korĵenkov 2005 and discussion below),
- Authoritative **grammars** and **dictionaries** written in Esperanto (Kalocsay & Waringhien 1985, Duc Goninaz 2002, Wennergren 2005), as well as bilingual dictionaries and learning resources (Wells 1989),
- A technical **terminology** documented by associations of professional terminologists,

covering many areas of science and technology (Blanke 1989).

4. Elaboration (unofficial sources of a self-regulating, emergent language norm):
- A wide range of **written genres** (including translated and original texts): anthologies, dictionaries, encyclopedias, magazines, novels, pamphlets, periodicals, web pages of international and national associations, as well as creative texts written for performance (song, theatre, poetry, etc.) (Auld 1984). The oldest text types in Esperanto are journals and magazines: the list includes some longstanding titles, such as current affairs (*Heroldo de Esperanto, Monato*), state-backed propaganda (*El popola Ĉinio* 'Out of People's China'), special interest journals (*Literatura Foiro, Scienco kaj Kulturo*), etc. There exist traditional libraries and dedicated text collections, increasingly made available on the internet in the form of digital text archives,[2]
- A range of **spoken genres**, including informal interactions and formal speeches at annual international congresses, conversations among co-workers in associations such as the *Universala Esperanto Asocio* and at local meetings. The spoken language is also available as recordings (speeches, readings, song etc.) and live radio programmes (broadcast in Europe notably by Radio Vatican and Radio Poland),
- An expanding range of **hybrid genres**, i.e. texts which exist in electronic form and are essentially interactive. This includes thousands of Wikipedia articles and other reference material, online courses, software applications adapted to Esperanto, individual webpages, blogs, videos and podcasts (Fettes 1997).

Yet despite this extensive social infrastructure, it has to be recognised that Esperantism has social characteristics which set it quite apart from most other language communities. One of the unique features of most neutral planned languages is that there is, in Gardner & Lambert's (1972) terms, hardly any 'instrumental' or utilitarian reason to learn them: there is no community of native speakers to look to, or mass-media to follow; neither is there any legal code, territory, state authority or state-backed educational system to defend or promote the language. But there is nevertheless a target language community: the highly distinctive community of Esperanto activists. It could be argued therefore that the main motivation for learning Esperanto is 'integrational', that is to say one actively seeks out to learn the language for cultural, ideological or psychological reasons, rather than financial, geographic, professional or other often coercive reasons. This view is supported by sociological surveys that have been carried out on the Esperanto movement (Forster 1982, Edwards 1994, Stocker 1996). These studies have pointed out that the typical Esperantist is relatively anti-conformist, well-educated, speaks several languages, often declares contrarian or left-leaning values such as 'internationalism', 'humanitarianism', 'green politics' and has sympathy for issues such as minority and regional language rights (see Fiedler 2006, Yaguello 2006, Okrent 2009 for some of the psychological as well as social reasons why people become Esperantists).

It is perhaps also worth adding another of the main reasons why some people

[2] For details of linguistic research on a large-scale corpus of Esperanto texts see the *Tekstaro de Esperanto* 'Corpus of Esperanto' webpage: http://tekstaro.com/. For examples, see analysis in section 5.

choose to learn Esperanto: the fact that there is *no* sizeable L1 speech community. As has often been observed, L2 users who not see themselves as very proficient in that language often prefer to deal with other L2 users, since this involves a degree of equality that is implicitly absent when L1 speakers are involved (Corsetti 1996). The propagandists of Esperanto have often argued that the great advantage of a neutral IAL as opposed to a national language is that in principle the IAL does not belong to any one speech community. However, the idea that Esperanto has no target language community is likely to be increasingly challenged if there is an increase in the proportion or status of L1 'native' speakers (Comrie 1996). Similarly, the fact that Esperanto has been developing an autonomous culture (Fiedler 2006), and with that a set of emergent norms and phraseology (and therefore an emergent elite of 'competent, proficient' speakers), may also pose a threat to the ideal of an egalitarian speech community which belongs to no one group of people (Fettes 1996).

Although the motivations and social profile of Esperantists have changed over the years, the Esperanto movement still represents one of the most unusual speech communities known to linguistics. Yet the heterogeneity of this community raises difficulties from the point of view of language planning: how can we recognise language norms and how are norms enforced in a largely uncontrolled speech community? How do we define a 'competent' speaker of Esperanto? Who among the competent speakers is qualified to judge the other users of the language? There are many questions of this sort, and it is beyond the scope of this paper to answer them all. Nevertheless, I do intend to address the particular question of whether there is an emergent norm in Esperanto (as claimed by Fiedler 2006). In the following section, I suggest a number of ways in which this norm has evolved.

3. The evolving norms of Esperanto before 1905

Textbooks and encyclopedias often state that Esperanto was 'created' in 1887. This account assumes that, once published, the language was a *fait accompli:* a fully developed system with an agreed set of rules about grammar, vocabulary and, at least implicitly, a functioning phonology and phraseology. The following quote by the Esperanto translator and linguist, Claude Piron, tells a different story. Piron argues that the development of the language was the result of a complex series of experiments and interactions, which he describes as follows:

> « L'espéranto est une langue jeune, née de la rencontre entre un intense désir de communication transculturelle au sein d'une fraction de la population du monde, et une proposition linguistique élaborée par un jeune polyglotte qui avait mis au point son projet en composant des poèmes et en traduisant des textes de différentes littératures. Adopté par des personnes d'origines ethniques et sociales dissemblables, ce projet, en servant à la communication, s'est naturellement transformé en une langue vivante par **un processus collectif, anonyme, largement inconscient, fait d'une série d'interactions et de réajustements**

mutuels... » (Claude Piron 1986, p11, my emphasis).[3]

As told by Piron, the creation of Esperanto was a 'co-production', in which the language scheme became a living language through a process of interactive language creation. Some have claimed (Lindstedt 2009 among others) that this process has more in common with the unplanned development of contact languages such as pidgins and creoles than the more conscious processes of language design or language planning.

It is now known that Zamenhof's *Internacia Lingvo* 'International Language' was preceded by several alternative versions of the language, and that Zamenhof had been experimenting on these for many decades, translating the classics and composing other texts, before finally settling on a (tentative) set of proposals. This period corresponds to the well-known process of 'selection' in language standardisation (Haugen 1972), and indeed Zamenhof (1936) later discussed the often subjective ways in which he had toyed with and ultimately selected one form from a variety of potential alternatives. The actual form of the language when published was therefore deceptively concise: *Internacia Lingvo* (later known as *Unua Libro* 'first book') was published anonymously by 'Doktoro Esperanto' (Zamenhof 1887) first in Russian and then in English, French, German, Hebrew and Polish. The book consisted of 40 pages, including:

- an introductory preface (28 pages, including example texts such as a prayer, excerpts from the Bible, original and translated poems and a sample letter),
- a set of 'sixteen rules' (6 pages, including 13 principles of grammar and word-formation, 2 statements on the phonetic principle of the alphabet and word accent and 1 statement allowing for the free adoption of 'foreign' i.e. international words),
- a set of 8 promissory notes (in effect, a petition promising to learn the language "if 10,000,000 people sign the same promise"),
- a bilingual glossary (Esperanto-Russian in the first edition, including 917 word roots, with 40 other roots being used in the rest of the booklet).

As pointed out by Wells (1989), the well-known 'sixteen rules' were incomplete, ambiguous and made many implicit assumptions. Rules 1 and 2 give a flavour of this:

> 1. There is no indefinite, and only one definite, article, *la*, for all genders, numbers, and cases.
> 2. Substantives are formed by adding *o* to the root. For the plural, the letter *j* must be added to the singular. There are two cases: the nominative and the objective (accusative). The root with the added *o* is the nominative, the objective adds an *n* after the *o*. Other cases are formed by prepositions;...

[3] "Esperanto is a young language, originating from the encounter between an intense desire among a fragment of the world's population for crosscultural communication and a language proposal drafted by a young polyglot, who had put his project together by composing poems and translating texts from different literatures. Adopted by people of differing ethnic and social backgrounds, the project began as a tool for communication and was transformed naturally into a living language by a process which was **collective, anonymous, largely subconsious, and made up of a series of mutual interactions and readjustments**..." (Claude Piron 1986, 11, my translation)

Statements such as these provided a useful summary of the morphological and lexical classes of the language (articles, nouns, case, etc.). But they gave virtually no guidance on usage. For example, the French version of Rule 1 had the extra remark (not present in the English version) that "the use of the article is the same as that of German, French and other languages...". This might have been of help to someone who knew these languages, but it hardly amounted to a valid linguistic description, and it was framed entirely in the traditional terminology of another language.

It follows from this that once the scheme was published, the language had to go through a further period of development, during which it was learnt by a few early followers and then advertised to the wider community (corresponding to Haugen's process of 'acceptance') at the same time as being used for increasingly diverse forms of communication (the process of 'elaboration'). In the first few months, Zamenhof received postcards and letters (and also visitors, who wanted to practise the language in person), then in the following decades the growing number of Esperantists started to correspond with each other and to organise meetings and conferences. During this time, Zamenhof and his correspondents expanded the stylistic range of the language, at first in the form of a monthly magazine *La Esperantisto* (Zamenhof 1889) and then in anthologies of poetry and prose such as *Fundamenta Krestomatio* (Zamenhof 1903).

The final stage in the development of a standard corresponds to Haugen's process of 'codification'. This is usually understood as the description and prescription of an established language norm. For most languages, this process is fraught with political difficulties. In the case of Esperanto the problem was compounded by the fact that many early Esperantists were language inventors and reformers, who often had their own particular preferences about one or another aspect of the language. Before 1905, Zamenhof had not objected to proposed changes, and in 1894 he had even asked the readers of *La Esperantisto* to vote on his own proposals for reform (known as *Reformita Esperanto*). These were intended to meet criticisms that had been made by Esperantists and outside observers over small issues such as the accented letters, or major issues such as the accusative. But in the event these reforms were rejected. The only area of consensus was on new lexical items, and so Zamenhof published a dictionary in 1894, *Universala Vortaro* ('Universal Dictionary') which contained the original 957 word roots plus 1740 new items, mostly derived from the emerging literature. It is significant that at this stage Zamenhof and his correspondents were not only developing a normalised vocabulary, they were also establishing the underlying grammatical principles of the language. The following statement from the *Universala Vortaro* gives an example of this:

> « Ĉion, kio estas skribita en la lingvo internacia Esperanto, oni povas kompreni kun helpo de tiu ĉi vortaro. Vortoj, kiuj formas kune unu ideon, estas skribataj kune, sed dividataj unu de la alia per streketo, tiel ekzemple la vorto « frat'in'o », prezentante unu ideon, estas kunmetita el tri vortoj, el kiuj ĉiun oni devas serĉi aparte. » (Zamenhof, 1894, p1)
>
> "Everything written in the international language Esperanto can be translated

> by means of this vocabulary. If several words are required to express one idea, they must be written in one but, separated by commas; e. g. « frat'in'o » though one idea, is yet composed of three words, which must be looked for separately in the vocabulary." (Zamenhof, 1894, p1 [this is the original English text]).

In other words, the Esperanto word *fratino* 'sister' is a compound made up of three roots: *frat-* 'brother', *-in-* 'ending of feminine words', and *-o* 'ending of nouns' (these are the glosses given in English in the *Universala Vortaro*). It has been inferred from this that Esperanto's morphology is 'agglutinative', a particular type of word formation system in which each word (lexeme) can be broken down into component roots (morphemes) which are themselves treated as a separate lexical entities, and thus separate entries in the dictionary. The system is simple to learn and highly productive, but it has led to theoretical problems (c.f. Gledhill 2000, and below).

However, Zamenhof's early attempts to codify the language could not prevent growing conflict from within the movement, notably in the form of counter-proposals and further attempts at reform. By 1905, the debates had become so divisive that he asked the first world congress of Esperantists at Boulogne-sur-Mer to vote on a stable form of the language, known as the *Fundamento de Esperanto* ('Foundation of Esperanto'). The text contained the original grammar from the *Unua Libro* (Zamenhof 1887), the extended vocabulary of the *Universala Vortaro* and a collection of exercises. At the same congress, Zamenhof and his supporters also proposed the creation of the *Lingva Komitato* ('Language Committee'), which later led to the creation of a language academy.

The declaration eventually voted by the congress included the following statement from the *Fundamento*:

> « Por ke lingvo internacia povu bone kaj regule progresadi kaj por ke ĝi havu plenan certecon, ke ĝi neniam disfalos kaj ia facilanima paŝo de ĝiaj amikoj estontaj ne detruos la laborojn de ĝiaj amikoj estintaj, ― estas plej necesa antaŭ ĉio unu kondiĉo: la ekzistado de klare difinita, neniam tuŝebla kaj neniam ŝanĝebla Fundamento de la lingvo. » (Zamenhof 1905, p1)

> "In order for the international language to be able make good and regular progress, and to be entirely certain that it will not fall to pieces, or that its friends in the future will not by some careless mistake destroy the work of their friends from the past, – there is one absolutely necessary condition: the existence of a clearly defined, untouchable and unchangeable Foundation of the language." (Zamenhof 1905, p1 [my translation].)

This declaration caused an immediate split in the movement. Some reformers and their supporters elected to abandon Esperantism, most notably Louis Couturat and Léopold Leau, who promoted an alternative language scheme, called *Ido* 'offspring' (Couturat & Leau 1908). This was not the only or the last example of schisms among Esperantists (see Smith 2011 for a discussion of the various counter-proposals). But in retrospect, the vote on the *Fundamento* is perceived by many observers as a turning point in the history of Esperanto: it produced a clear normative statement about the

core features of the language, it established a democratic precedent for the community to follow in later years, and it had the immediate effect of causing the most vocal critics of the language to leave.

4. The evolving norms of Esperanto after 1905

As we shall see, the story of the codification of Esperanto is about how the language eventually freed itself not only from its 'parent' languages but also from other competing language schemes, at the same time as extricating itself from traditional linguistic descriptions (a process that has also been observed in modern languages such as English and French in relation to the old Latin grammars). As mentioned above, the Esperantists who were elected to join the *Lingva Komitato* had an important role to play in this process. Soon after its creation, the *Lingva Komitato* set up a subgroup which was to become the *Akademio de Esperanto* ('Academy of Esperanto', 1908). The following table sets out a summary of their recommendations since 1967:

Year	**Title**	**Topic**
1967 (1989)	*Pri pasivaj participoj*	Morphology / Syntax: difficulties with the present passive participle *-at-* and past passive *-it-*.
1967	*Enketo kaj decidoj pri la Vortfarado*	Morphology: enquiry into and recommendations on word formation (difficulties with the interpretation of root words).
1971	*Pri la refleksivo*	Syntax: difficulties with the mediopassive root *-iĝ-* and the reflexive pronoun *si*.
1974	*Pri la vorto "po"*	Syntax: difficulties with the preposition *po*.
1974, 1985, 1989, 2003	*Pri landnomoj*	Terminology / Toponyms: use of the neologism *-i-* versus *-uj-* to derive the names of countries.
1982	*Pri niaj alfabeto kaj ortografio*	Alphabet / Orthography: difficulties with transliteration, pronunciation.
1986	*Esprimo de la aganto en komplementa funkcio*	Syntax: which preposition to use when a complement expresses the role of agent (the neologism *far* versus the original preposition *de*).
1989	*Rekomendoj pri propraj nomoj*	Proper nouns / Anthroponyms: difficulties with transliteration, pronunciation, gender...
1990	*Pri derivaĵoj de komun/a: komunumo, komunaĵo, komuno*	Morphology: difficulties with the derivation and interpretation of root words, notably *komun/a* 'common' *komuno / komunumo* 'community', etc.
2007	*Pri apartaj teknikaj bezonoj rilate al niaj alfabeto kaj ortografio*	Alphabet and Orthography: difficulties arising from advances in technology.

This list gives us a useful picture of the problems and hesitations encountered by Esperanto speakers over the years, and tells us much about the incremental changes that have taken place in the evolution of the language. Unlike the *Académie de la langue française,* the *Akademio de Esperanto* has traditionally included a number of academic linguists and language professionals over the years (more recently: Probal Daŝgupta, Michel Duc Goninaz, Paul Gubbins, Katalin Kováts, Jouko Lindstedt, Tsvi Sadan, John Wells, Bertilo Wennergren...). It is perhaps for this reason that the recommendations of the *Akademio* have generally adopted a descriptive rather than a prescriptive approach: when a particular form comes to replace an older form in attested language use, the *Akademio* is generally prepared to admit the new form (and if the older form is in the original *Fundamento,* the new form is admitted as an 'alternative' to the older form.)

The topics discussed by the *Akademio* fall into three main types: 1) Decisions about a novel form which may replace an existing form in the *Fundamento*. For example, the decision to accept the form *-i-* for any country name which is derived from the name of a people, as in *franco* 'French (person)' > *Francio* 'France' (a usage which has mostly replaced the root *-uj-* originally in the *Fundamento,* as in *Francujo* 'France'). 2) Attempts to resolve grammatical problems, often involving issues which date back to the original language scheme (hesitations about the use and interpretation of passive participles, the preposition *po,* the reflexive pronoun *si,* etc.). 3) Recommendations about neologisms and alternative lexical forms, i.e. where two or more alternative forms are structurally possible or attested in different sources (*aŭtenta vs. aŭtentika* 'authentic', *meĥanisto vs. mekanikisto* 'mechanic, mechanician', *spontana vs. spontanea* 'spontaneous', *redaktisto vs. redaktoro* 'editor'), or where a new notion may have several competing names (*komputilo vs. komputoro* 'computer', *televido vs. televizio* 'television', and more recently *datao vs. dateno, datumo, donitaĵo,* all of which have been proposed for 'data').

At the same time as considering problems of usage, the *Akademio* was mandated to produce approved vocabulary lists and to develop technical terminology. When official words are cited in dictionaries, they are given a number, to indicate which of the Academy's supplements they have come from and to show their status in relation to 'fundamental' vocabulary (all items published in the 1905 *Fundamento*). This attention to the official status of lexical items is a particularity of Esperanto lexicography, and stands in contrast to other languages, where etymology or date of first attestation is often seen as more relevant. To date the Academy has published 9 official supplements, expanding the official vocabulary to a current total of 4758 official root words and grammatical endings (this includes the fundamental vocabulary of 2768 words, plus 2185 roots codified by the academy). These totals include several productive roots used in the sciences (including roots such as *-som-* 'non-plastid organelle', as in *kromasomo* 'chromasome', *ribosomo* 'ribosome', etc.), but the Academy recognises many more attested terms. The Akademio also records the changes in meaning of certain roots over the past century. For example, some words have undergone semantic generalisation (*klopodi* 'to endeavour', > 'to try', *ŝati* 'to esteem' > 'to like'), whereas others have become more specialised or have

changed register (*drogo* > 'chemical' > 'narcotic', *plena* 'complete' > 'full').

As with most other academies, the work of the *Akademio de Esperanto* is often hidden from view: for most language users, it is the reference dictionary which is seen as the established source of legitimate usage, both for lexical and grammatical issues. In recent years the most authoritative monolingual dictionary in Esperanto has been the *Plena Ilustrita Vortaro* 'Comprehensive Illustrated Dictionary', known as PIV (Duc Goninaz 2002). PIV includes the fundamental roots and official words recommended by the *Akademio*, but the dictionary makes use of a much greater corpus of sources. In addition, the agglutinative nature of the language means that many thousands of compound words are formed on the basis of approximately thirty highly productive 'root' morphemes (such as *-ar-* 'collection of', *-ej-* 'specific place for', *-ist-* 'professional', etc.). To give just one example, the root *-il-* 'tool' is used to form composed words, such as *flugilo* 'flying tool, wing', *komunikilo* 'communication tool, medium', *ligilo* 'linking tool, hyperlink', *montrilo* 'showing tool, pointer', '*skribilo* 'writing tool, pen', *trancĥilo* 'cutting tool, knife', etc. This kind of word formation system accounts for the large number of 'entries' in the dictionary compared with 'root words.' The current edition of PIV lists over 17,000 simple 'root' words and 47,000 entries for compound words. These figures are comparable with reference dictionaries in other languages (the *Petit Robert,* for example contains 60,000 entries in French). A final point regarding vocabulary is that the pioneer of terminology studies Eugen Wüster developed his approach to terminology in parallel with his activities as an Esperantist: it is no accident that there is a very strong tradition of specialised terminological work in Esperanto (most notably from 1987 with the creation of the *Terminologia Esperanto-Centro*) (Haupenthal 1978, Blanke 1989, 2008).

So far we have seen how institutions such as the *Akademio* have worked to defend the standard established in 1905 or to gradually reflect the changing norms of Esperanto. But the activities of these authorities can have little effect on language change. For the most part, language change in Esperanto has been initiated by individual authors, for example the Czech, Karolo Pič who was responsible for many inventive and provocative word forms which have effectively expanded the stylistic resources of the language. From time to time, there have also been broader movements to introduce reform, largely as part of a more general social trend. A recent instance of this involves the debate on sexist or gender-biased language. In fundamental Esperanto (Zamenhof 1887, 1905), as we have seen, the traditional way of feminising a word was to use the root morpheme *-in-* to form a compund with an animate word: *frato* > *fratino* 'brother > sister', *kuracisto* > *kuracistino* 'doctor > female doctor', *bovo* > *bovino*, 'bull > cow', *kato* > *katino* 'cat > she-cat', and so on. If a specifically male noun was required, the root word *vir-* 'man' was used as a prefix (> *virbovo* 'bull', *virkato* 'tomcat' etc.). However, there was much hesitation about this usage and words for male humans were considered redundant (**virfrato*) or unnecessary (?*virkuracisto*). But in a recent response to feminist criticism of the language, some Esperantists have proposed that root words should as far as possible be considered to have no inherent gender. Some have even proposed the morpheme *-iĉ-* in order to create compound words such as *boviĉo / bovino* 'bull / cow', *katiĉo /*

katino ‘tomcat / she-cat’ and so on. The authoritative grammars[4] are prepared to accept a distinction of this kind for words which do not have an inherent masculine meaning, such as *kuracisto* > *kuracistiĉo* / *kuracistino* ‘male / female doctor’. But there is more hesitation about using the morpheme with pairs of words which have until now conventionally stood in contrast to feminine forms, such as *edzo / edzino* ‘husband / wife’. Similarly, few users would accept the use of the new morpheme for an inherently masculine form such as **fratiĉo,* although there is no reason why the word *frat-* should not refer to ‘sibling’, and it is legitimate to ask why a word such as ‘sister’ has to be derived from ‘brother’ and not the other way round (and similarly *patrino* ‘mother’, *onkino* ‘aunt’ etc. all derived from default masculine roots).

The proposal to adopt the *-iĉ-* morpheme can thus be seen partly as a response to the logical loopholes in the original design of the language, but also as a more general trend among Esperantists to respond to long-standing accusations of structural sexism. It is interesting to note however how difficult it is to propose reforms of this sort, regardless of how well founded they may be: the discourse practices and expectations of many Esperantists are now too firmly established for such reforms to be implemented immediately.

I mentioned above that the language experts and authorities have come to reevaluate their analysis of the morphological structure of the language. This new conception of Esperanto’s morphological system is part of a wider movement to ‘free’ the language from the analytical traditions of other languages. In this respect, it is important to mention the latest comprehensive grammar to be published, Wennergren’s (2005) *Plena Manlibro de Esperanta Gramatiko* (‘Comprehensive Handbook of Esperanto Grammar’, known as PMEG). PMEG diverges quite sharply from previous attempts to describe the language, most notably Kalocsay & Waringhien’s (1980, 1985) *Plena Analiza Gramatiko* (‘Comprehensive Analytical Grammar’, known as PAG). PAG was a very detailed, descriptive grammar, and its terminology owed much to mainstream linguistics analysis (such as *adjekto* ‘adjunct’ *konjunkto* ‘conjunct’, *eksplika kompletivo* ‘explicit completive’, etc.). Much space in PAG was given over to the problems debated by the *Akademio*, and its aim was to provide a synthesis of some of the most complex debates (most notoriously the nature of word roots). The aim of PMEG on the other hand is to provide practical guidance for ordinary users of Esperanto. In order to do this, PMEG has invented a vocabulary of its own for the basic word forms of the language:

1. Finaĵovortoj	‘words requiring a grammatical ending’ (such as *inklud-, -ind-* etc.),
2. Rolvortetoj	‘function words not requiring a grammatical ending’ (such as the article *la,* the preposition *de,* etc.)
3. Finaĵoj	‘grammatical endings’ (such as *-a, -i, -oj,* etc.)

PMEG also proposes new forms for traditional terms such as *akuzativo* ‘accusative’, relpaced by *rolfinaĵo -n* ‘functional ending -n’, *substantivo* ‘noun’ replaced by *o-vorto* ‘o-word’, and so on. There are disadvantages in using this kind of

[4] Notably *Plena Manlibro de Esperanta Gramatiko* (Wennergren 2005, discussed below).

terminology. But I would suggest that by using novel compounds in order to explain grammatical terms, PMEG reflects a new approach to how Esperanto is perceived by its speech community. It is perhaps no accident that the author of PMEG is a member of the *Akademio* and a 'corpus linguist': as such his is familiar with the contemporary tendency for linguistic metalanguage to be couched in ordinary language. This notion owes much to empirical linguists, such as J.R. Firth, a follower of Wittgenstein who coined the phrase 'language turned in on itself'.

5. The emerging phraseology of Esperanto

In the previous sections, we have seen how various authorities (Zamenhof, the *Akademio*, dictionaries such as PIV, grammars such as PAG and PMEG) have gradually developed the 'explicit' norms of Esperanto, that is to say the visible features of language which are often identified in reference works and discussed by language experts and other commentators. In this section, I examine the notion of an 'implicit' norm, that is to say the underlying regularities of expression which are not available to conscious introspection. In particular, I focus here on one of the most interesting manifestations of implicit norms in language: 'phraseology'.

For many linguists, phraseology refers to the study of idiomatic expressions, proverbs, conversational formulae and other more or less idiosyncratic constructions. Zamenhof himself was aware of the need to provide a collection of phrases which would serve as a structural and metaphorical framework for the language, as can be seen in his *Proverbaro Esperanta* 'Collection of Esperanto Proverbs' (Zamenhof 1910) which listed 2630 proverbs and other expressions in Esperanto, French, German, Polish and Russian. Most Esperantists are familiar with a scattering of these formulae, as pointed out by Fiedler (2007). One famous example, still commonly used today, is the rhyming phrase: *Ne ŝovu la nazon en fremdan vazon* 'Don't stick your nose in someone else's business (lit. vase).'

In recent years, the use of computer-held corpora has revolutionised the ways in which grammarians and lexicologists view their discipline, especially those with a more empirical or functionalist orientation (Halliday 1985, Sinclair 1991, Tucker 2007 inter alia). For these linguists, phraseology corresponds more broadly to the particular speech style of an individual, a type of discourse or a group of people. This happens to be how the word is used in the general language, in such phrases as: *the phraseology of + the Far Right, Jane Austen, administrative French,* as well as in other languages: *die Phraseologie von Marx, la fraseología de los textos jurídicos, la phraséologie aéronautique,* etc. Thus whereas traditional phraseology is preoccupied with the classification of different phraseological units (idioms, routine formulae, etc...), the approach adopted by these linguists is to consider that any meaningful sequence of words which can be observed in a particular text is an instance of a more general 'lexicogrammatical pattern'. The notion of the lexicogrammar is central to Halliday's theory of a language:

> There is in every language a level of organization – a single level – which is

> referred to in everyday speech as the 'wording'; technically it is a lexicogrammar, the combination of grammar and vocabulary. (…) The point is that grammar and vocabulary are not two different things; they are the same thing seen by different observers. (Halliday 1992:63)

In practice, this means that every lexical item in a language is associated with a set of more or less pre-constructed grammatical structures, and each grammatical structure is associated with a relatively predictable paradigm of lexical items. The lexicogrammar approach also suggests that the regularities of speech which can be observed in a representative corpus of texts are ubiquitous (shared across the speech community) and subconscious: they cannot usually be found intuitively by asking language users what they think, but have to be observed in a representative corpus of texts. The observation of lexicogrammatical patterns has become a standard methodology in the fields of lexicography and corpus-based descriptive linguistics (see accounts in Sinclair 1991, Hausmann & Blumenthal 2006, Béjoint 2007, Tucker 2007, *inter alia*).

It is my purpose here to demonstrate that the phraseology and lexicogrammatical patterns of Esperanto are consistent and predictable, and can thus serve as a useful linguistic marker of group identity, in a similar way that pronunciation, orthography or lexis serve this purpose at other levels of the language system. The hypothesis I raise in this section is whether the phraseology of Esperanto has remained consistent or has become fragmented, and whether its phraseology has developed in the same way that it would in natural languages. In order to answer these questions, in the following two subsections I examine two specific examples of phraseology in Esperanto: light verb constructions (LVCs) build around the light verb *fari* 'to do, to make' and LVCs build around *havi* 'to have'. It is important to note here that all of the following examples are attested, and come from a large online corpus of Esperanto texts (the *Tekstaro de Esperanto*: 4,675,412 words of translated and original works).[5]

5.1 The phraseology of *fari* (to do, to make)

Has Esperanto inherited its underlying lexicogrammatical patterns from one language, or does Esperanto's phraseology reflect the influence of several languages? As has been stated elsewhere (Wells 1989, Gledhill 2000), over 70% of Esperanto's lexical morphemes were taken from the Romance languages (with the rest coming form the other major western European languages), but this does not mean that the phraseological patterns that are associated with these items have also been carried over into Esperanto. One particular example involves 'light verb constructions' (LVC), that is to say extended verb phrases in which a verb expresses a generic 'light' meaning ('to do, to have, to make' etc.), while a noun specifies the precise meaning of the phrase as a whole ('to do some work, to have a bath, to make sense', etc.). The

[5] Available at: www.tekstaro.com. The creation of the corpus was coordinated by the *Esperantic Studies Foundation* and the linguist Bertil Wennergren. The present author was also involved in this project.

distribution of these phrases is often different in Esperanto and its donor languages. Because of its agglutinating morphology, for example, many constructions which involve light verbs in English are expressed by a single verb or a compound verbal root in Esperanto, as in *sin amuzi / amuziĝi* 'have fun', *amiki / amikiĝi* 'make friends' *amori / seksumi* 'make love / have sex', *pravi* 'be right, make sense', etc. But Esperanto does make use of LVCs in other contexts, and in many cases the light verbs used (*fari* 'to do', *havi* 'to have', *doni* 'to give' etc.) do not correspond to those in other languages:

fari baton	'(to do / make) give a slap' (cf. Fr. *donner un coup*)
fari demandon	'(to do / make) ask a question' (cf. Fr *poser une question*)
fari omaĝon	'(to do / make) pay homage (to)' (cf. Fr *rendre hommage*)
fari viziton	'(to do / make) pay a visit (to)' (cf. Fr. *rendre visite*)

Although *fari* appears to be the default verb in many contexts, there are alternatives, which often correspond to one or another donor language:

doni atenton	'(to give) to pay attention' (cf. Fr. *faire attention, prêter attention*)
jeti rigardon	'(to throw) to have a look' (cf. Fr. *jeter un coup d'oeil*)
tiri atenton	'(to pull) to attract attention' (cf. Fr. *tirer l'attention*)
tiri konkludon	'(to pull) to come to a conclusion' (cf. Fr. *titer une conclusion*)

The verb *fari* is however by far the most productive light verb in Esperanto, as can be seen in the following selection of examples. Note that in the following analysis, I have underlined the light verb construction in English (but I have only given a minimal translation).

It is a general principle of the lexicogrammar approach adopted here that each particular construction in the language has its own unique lexical and grammatical particularities, and its own particular discourse function or meaning. Thus examples 1-4 below are part of a more general pattern, in which the verb *fari* is ditransitive (it takes two complements) and the indirect complement is positioned immediately after the verb (as in *al mi* 'to me', *al li* 'to him(self) etc.). Notice that in terms of meaning, this construction generally expresses a cognitive or communicative process (and it is notable that in English the verb *fari* is not always translated by its usual equivalents, 'do' or 'make'):

1. *"Mi el nenio **faras** al mi **ĉagrenon**!" diris la ombro, "kaj tial mi grasiĝas..."*
   'Out of nothing I cause myself grief'...
2. *Via aparta letereto, en kiu vi komunikis al mi la peton de unu certa persono interrilatigi min kun li, **faris** al mi grandan **impreson**.*
   'Your separate letter... made a great impression on me'
3. *Antaŭ tri tagoj mi vizitis vian kuzon kaj mia vizito **faris** al li **plezuron**.*
   '... my visit gave him pleasure'
4. *mi vidis grandan nigran spiriton, kiu etendis la brakojn kaj **faris** al mi **signon** ne iri plu.*
   'I saw a great black spirit, which spread its arms and made a sign for me not to go on'

There is however a second lexicogrammatical pattern associated with *fari, w*hich involves a material (dynamic or physical) process (examples 5-8). Here *fari* is monotransitive, and does not involve an indirect pronoun. Again, the translations often involve different light verbs in English:

5. *Peano **faras** pioniran **esploron** por matematika lingvistiko;*
   'Peano carried out pioneering research for mathematical linguistics'
6. *William AULD **faris** grandan **laboron**, kolektinte, krom la tri supremenciitajn poemarojn, ankaŭ multajn aliajn poemojn,*
   '...William Auld did a great job...'
7. *Li kaj lia amiko jam priparolis ĉion kaj jam delonge **faris** en la konversacio longajn **paŭzojn**, nur frapetante unu la alian je la femuroj kaj dirante...*
   'He and his friend had already talked about everything and had taken long pauses'
8. *Ie tie supre, se kredi la agoniajn vortojn de ŝippereulo, kiu **faris** la timindan tralandan **vojaĝon**, kaj se veron atestis la flakono da oraj grajnoj...*
   '(to make) undertook a frightening cross-country journey'

Let us now examine the phraseology of one construction: *fari demandon* 'to ask a question' and compare it with some related constructions in Esperanto. The aim of this analysis is to show that Esperanto has inherited a range of competing constructions from various languages, but has developed a consistent but also contrastive context of use for each one. In the *Tekstaro,* it is possible to find four different light verbs for this construction:

- ***fari** demandon* 'to ask question' (equivalent to Spanish *hacer una pregunta,* and Italian *facere una domanda*),
- ***starigi** demandon* ('to pose a question' equivalent to French *poser une question*, and German *eine Frage stellen*),
- ***meti** demandon* ('to put a question', English *raise a question*),
- ***levi** demandon* ('to raise a question', French *soulever une question*).

Since these alternatives originate from a variety of languages, it might be thought that they are equivalents and thus exist in free distribution. However, an analysis of the corpus suggests that not only do these LVCs express different nuances (suggested by the light verbs which they are constructed with), they are also used in very consistent and distinctive lexicogrammatical contexts. To demonstrate this, let us compare the typical contexts of the simple verb *demandi* 'to ask' with its 'equivalent' light verb constructions. As is often the case (see Gledhill 2008 for further discussion of this), the simple verb can be is used in a variety of lexicogrammatical contexts, although most often it occurs in close proximity to a question in the interrogative (often formed by the linking word *ĉu* or question words such as *kiam* 'when'):

9. *- Ĉu vi kontraŭas, ke mi rigardu vian registrolibron? - **demandis** Holmso.*
   (indirect speech '... - asked Holmes')
10. *La dua parolanto ne **demandas**, ĉu la unua estas barono Brambeus.*
    (complement clause 'the second speaker asked if...')
11. *Enirinte en karavanejon, li tuj **demandis**, **kiam** komenciĝos la sklavovendado...*

(complement clause ‘he immediately asked when will... begin’)

12. *ŝi ja **demandis ilin pri** la nomo de la reĝidino;*
(indirect complement ‘she asked them about the princess’s name...’)

If we compare these examples with *fari demandon,* we can see that the construction is almost exclusively used in the context of complement clauses, formed by the subordinator *ĉu* ‘whether’ (examples 13-16). This usage is attested very early on in Esperanto, in the speeches of Zamenhof (here cited in example 13). It is notable that in these contexts, *fari demandon* often has the meaning ‘ask permission’:

13. *en la fino de l’ kongreso mi **faros** la **demandon**, ĉu vi akceptas mian manieron de konduko*
(‘at the end of the congress, I shall ask (the question) whether you accept the way I have conducted myself...’)
14. *La Princo ekprenis la skatolon, kaj malfermis la buŝon por **fari** alian **demandon** ; sed, antaŭ ol li povis eligi unu solan vorton […].*
(‘the Prince... opened his mouth to ask another question...’)
15. *"Sed, Baptopatrino, mi petas, ĉu mi povos **fari** al vi unu **demandon**?"*
(‘can I ask you one question...’)
16. *En unu kunveno esperantista iu **faris** la **demandon**, ĉu oni devas diri: “mi amas vin”, aŭ “mi vin amas”, […].*
(‘someone asked the question whether...’)

In contrast, the LVC *starigi demandon* ‘to pose a question’ (literally ‘to make stand’) involves a specific mental process, often in the sense of ‘issue, problem, or puzzle’. In these examples (17-20), it is notable that the noun *demandon* is usually post-modified by a preposition or a relative clause (a structure that is different to that of *fari demandon*):

17. *“Ĉu vi sentis timon?” mi diris, **starigante** la **demandon** kiun virinoj ofte **starigas**, kaj ricevante la respondon kiun preskaŭ ĉiam ili ricevas.*
(‘Did you feel afraid’ I said, asking the question that women often ask...’)
18. *Se li rapidis, tio **starigas** interesan **demandon**, kial li estis tiel rapidema […].*
(‘if he did hurry, that raises an interesting question..’)
19. *...nia lando estis la unua, kiu ankoraŭ en 1946 **starigis** la **demandon** pri malpermeso de produktado kaj uzo de atoma armilaro […].*
(‘our country is the first... to raise the issue of banning the manufacture and use of nuclear weapons...’)
20. *La romano ja **starigas demandojn** pri la universo kaj ties enloĝantoj, sugestas plurajn respondojn, sed lasas, ke la legantoj plu meditu.*
(‘the novel raises issues about the universe...’)

Examples 21-25 show that the construction *meti demandon* ‘put a question’ is often used in administrative or political contexts, where it refers to requests that are ‘submitted’ to a group to be decided collectively. In this pattern, we again get a structure which is different from *fari / starigi demandon.* This time, the noun is unmodified, and stands in close proximity to the verb, thus allowing the indirect

complement, which is longer and usually more complex, to be placed later on in the clause:

21. *Mi **metis demandon** al la grupo pri la aervorticoj kaŭzintaj la lastnovembran katastrofon de aviadilo proksime al la aerhaveno…*
    (‘I put a question to the group...’)
22. *Tio povas okazi ekz. tiamaniere, ke la komputilo **metas demandon** al la uzanto de la komputilo, kiam estiĝas dubaĵo, atendas respondon kaj daŭrigas la...*
    (‘...the computer submits a question to the user...’)
23. *En la pasinta numero de "Japana Budhano"(n-ro 298) s-ino KOBAYASI **metis** demandon, ĉu budhismo malpermesas sinmortigon aŭ ne.*
    (‘...in the last edition, Mrs Kobayasi sent in a question about whether Buddhism prohibits suicide...’)
24. *Mi **metis demandon** sur la reton antaŭ du monatoj, sed ricevis nur du respondojn.*
    (‘I put a question on the web two months ago...’)

Finally, examples 25-28 show that the LVC *levi demandon* refers to abstract arguments in which the question marks a turning point or transition in an ongoing text. As with *meti demandon*, the noun in *levi demandon* is less heavily modified than in the construction *starigi demandon*:

25. *Tiu ĉi scikono de sia tuto **levas** du **demandojn**.*
    (‘this awareness raises two issues’)
26. *Parenteze, oni povus **levi** la **demandon**: Kial SAT-anoj […] malsimpatias al modernaj artformoj?*
    (‘by the way, one might raise the following question:...’)
27. *Mi **levas demandon** pri administro de TTT-paĝoj, i.a. de UEA.*
    (‘I am raising the issue of how webpages are administered...’)
28. *Tiuj procesoj **levis** ŝlosilajn **demandojn**.*
    (‘These processes raise two key points’)

There are of course no clear rules about when to use *fari, starigi, meti* or *levi* (or other verbs) with the complement *demando(n),* any more than there are clear rules about similar constructions in English, French or other languages. It is interesting to note however that Esperanto has in effect ‘borrowed’ several light verb equivalents from each of its major donor languages (here English, French, German with possibly some influence from others) and has developed a specialised meaning for each. Thus overall, *fari demandon* tends to be used in rather more formal contexts, and its meaning is ‘ask permission’ (a usage regularly associated with complement clauses). The LVC *starigi demandon* is used to denote the formulation of intellectual, cognitive ‘problems’, which are often specified by grammatical post-modification. By contrast, *meti demandon* is used in communicative contexts in which a question is submitted to another person or group. *Levi demandon* is similar to *meti demandon,* but it is used as a rhetorical device as part of an on-going exposition.

5.1 The phraseology of *havi* (to have)

The function of *havi* 'to have' as a light verb in Esperanto is quite different to that of *fari* 'to do, make'. While *fari* is used in a range of light verb constructions which express a 'material' or a 'mental' process, *havi* is used in LVCs to express a 'relational' process, in which the noun expresses a degree of modality (obligation or possibility). This usage also exists in other languages (English and French, for example), although as we shall see below, the complicating factor is that for every modal verb such as *devi* 'to have to' in Esperanto there is a corresponding LVC *havi la devon* 'to have the duty'. I would argue that there is a consistent, contrast in meaning between these two forms. Thus in following examples (29-32) the verb *devi / devas* 'has to, must' expresses a dynamic 'actual' meaning, whereas *havi la devon* 'has the obligation to' expresses a stative 'virtual' meaning:

29. *Mi* ***devas*** *agi lerte...*
'... I must act carefully...'

30. *oni* ***devis*** *informi la parencojn en Beogrado...*
('someone had to inform the relatives in Belgrade...')

31. *la aliaj Membroŝtatoj* ***havas*** *la* ***devon*** *helpi kaj asisti ĝin per ĉiuj rimedoj je ilia dispono, laŭ Artikolo 51 de la Ĉarto de [...].*
'... other member states have the obligation to help...'

32. *Mi respondos, ke mi scias nenion, ke la afero same misteras al mi, kiel al ili, kaj ke mi preferus ne* ***havi*** *la* ***devon*** *ilin informi.*
('...I would prefer not to have the duty of informing them')

In examples 29 and 30, the main verb in the phrase is not *devas,* but *agi* or *informi*. The modal *devas* does not express a specific semantic process here; rather the subject of the predicate *mi* is seen as the agent of the following verb. By contrast, in 31 and 32, the subject is at the same time a subject of *havi* with the relational meaning of 'to have, to possess (as a characteristic)', as well as the implicit agent of the process expressed by the noun *helpon* 'help'. In these examples, the structures and meanings are basically the same in Esperanto and in English. But it is interesting to see how Esperanto has developed a series of modal expressions which not only parallels those of its 'donor' languages, but in fact goes beyond them. This can be seen in the following examples (33-36) of *ebli* 'to be able' and *havi la eblecon* 'to have the ability to', where *-ebl-* is an adjectival root derived from the latinate ending '-able', and on its own means 'capable, possible', while the root *-ec-* is used to derive nouns from adjectives and on its own means 'quality'. As mentioned in previous sections, Esperanto has the morphological capacity to derive verbs from nouns and vice versa very productively, so that in this case various alternative formulations are possible, such as *havi la eblecon* and *havi la eblon* 'have the ability' as well as simply *ebli* 'to be able to'. There is a subtle difference between these formulations: the single verb *ebli* is used to formulate an explicit evaluation of the following verbal process (just as though it were a modal verb such as *devi*), whereas the LVCs *havi la eblecon / havi la eblon* denote relational states and serve to define or characterise the subject rather than to express evaluation:

33. *Ne **eblas** trotaksi la rolon de ekonomiaj normoj.*
('it is not possible to overestimate the role of economic norms')
34. *Ni petas vin - se **eblas** - informi pri nia simpozio*
('We ask you, if (it is) possible, to inform (others) about our symposium')
35. *La gazetoj ne trovas necesa eniĝi en tion, pri kio ili skribas, kaj kontentiĝas nur per tio, ke ili **havas** la **eblon** regali la legantojn per ridinda novaĵo aŭ fari spritaĵon.*
('... magazines have the ability to delight their readers...')
36. *Dum lia vizito, preskaŭ 20 lokaj esperantistoj **havis** la **eblecon** renkontiĝi kaj paroli kun li dum vespermanĝo ĉe la korea restoracio NY Kom Tan.*
('...local Esperantists had the possibility of meeting up and talking...')

I will not repeat this analysis in the following examples, because in each case it can be seen that the simple verb construction expresses an 'actual', concrete variation of a more abstract, virtual relation expressed by a light verb construction of the form *have* + 'predicative noun':

37. *Sed mi ne **kapablas** forgesi ilin, fraŭlino Stepeltono, – mi diris.*
('I am not capable of forgetting them...')
38. *Kvankam la juna virino poste mortis, ŝi **havis la kapablon** forlasi sian tombon nokte kun la celo suĉi la sangon de aliaj personoj.*
('...she had the ability to leave her tomb at night...')
39. *Li **intencis** fine akiri la bienon, kaj li pretis uzi iun ajn ilon aŭ riski ion ajn tiucele.*
('he intended at last to acquire the property...')
40. *Mi **havas la intencon** prezenti al la Akademio projekton de regularo pri la tri jenaj punktoj : […].*
('I have the intention to presenting a project to the Academy...')

The final examples (41-44) are similar, but it is notable that the single verbs (here derived directly from the nouns *kuraĝo* 'courage' and *kutimo* 'custom, habit') do not have the same equivalents in English (or in the donor language, in this case French):

41. *La sekvan tagon la konsilisto estis tre gaja kaj neniu **kuraĝis** demandi lin pri la okazintaĵo de la nokto.*
(... no-one dared to ask him about the occurrence in the night...)
42. *[…] malgraŭ la diktaturo de la nuna prezidento Pervés Muŝarraf, apogita de Usono, ŝi **havis la kuraĝon** reveni al sia lando,...*
(... she had the courage to return to her country...)
43. *Kiam Kruko estas en kompanio de pluraj homoj, li **kutimas** demandi : – ĉu vi aŭdis pri la nova bordelo en la X-strato ?*
(...he tended to ask...)
44. *La amerikano ekzemple **havas la kutimon** meti tuj sub la nomo de la adresato la straton kaj numeron, […].*
(the American for example has the habit of putting the street and address...)

Examples such as these suggest that in languages like French and English there is no systematic contrast between 'simple' verb constructions and 'complex' light verb

constructions, or at best the distinction in these languages is expressed by lexical items which are morphologically unrelated ('dare', 'tend' in the above examples). In Esperanto, however, we appear to have found a very regular and productive paradigm, in which there is a simple verb (based often on a noun) and an equivalent light verb construction of the form *havi* + Predicative Noun. Since all the LVCs built around *havi* are used to express static, relational processes, it is not surprising that these constructions are often found in administrative, technical texts, in which relational processes such as 'to be' and 'to have' often have a more important role to play than in other text types, since they are typically used in the formulation of impersonal expressions.

6. Conclusion

In this paper, I have traced the development of both explicit and implicit norms in Esperanto. The explicit norms of Esperanto can be seen as the visible, tangible forms of expression which have been selected and codified by the various authorities of the language (the inventor Zamenhof, the *Akademio*, the reference dictionaries and standard grammars) as well as other influential participants (individual authors, pressure groups, etc.). I have argued here that the explicit norms of the language were not in fact formulated or even fully fleshed out in the original scheme of 1887: rather, the language itself has been (and still is) an ongoing 'work in progress', a co-production of many thousands of complex interactions and negotiations.

I have also attempted to identify some of the implicit norms of Esperanto, in an area of language use that has often been marginalised: phraseology. Implicit norms are by definition covert, tacit and uncontrolled, yet it has been shown in many other studies that these norms correspond to the ubiquitous and unconscious patterns of language use which underly all types of discourse, whether formal or informal, spoken or written. In the above analysis of an online corpus, I examined a particular instance of this in a family of expressions known as 'light verb constructions'.

While most Esperantists are familiar with phrases such as *fari demandon* '(to make) ask a question', *starigi demandon* '(to stand) pose a question' and so on, few language users (either L1 or L2) are able to state with any certainty the relative frequency or distribution of these expressions, let alone the lexicogrammatical co-texts or discourse contexts in which they are typically used. It might have been thought that Esperanto had borrowed these and other related constructions more or less at random from its 'donor' languages (often French, sometimes English and German) and now uses them indiscriminately, as though they were synonyms. The corpus analysis set out above shows that Esperanto has indeed borrowed a very eclectic mix of different light verb constructions from different languages, yet in each case each construction is used in highly consistent grammatical structures and with subtly distinctive meanings. This kind of evidence suggests that over the past century Esperantists have unconsciously developed a general phraseology, which one might compare with that of an 'accent' or 'speech style'. This observation appears to support Fiedler's claim (2006) that the language use of Esperantists is not merely held

together by a set of conscious rules and authoritative statements (the explicit norms of the language) but also by what she terms 'self-regulation': the constant intercourse of language users as they elaborate, incorporate and reproduce the language patterns of others. In Fiedler (2007), it was suggested that the formulation and reformulation of traditional forms of phraseology (idiomatic expressions and the like) give good evidence for the continued activity and vitality of a self-regulating speech community of Esperantists. In this paper, I have suggested that another kind of phraseology, the predictable yet also productive patterns of lexicogrammar, also provides strong evidence for the effective existence of an implicit form of language norm which, apparently, serves as a subtle badge of membership for this rather unique speech community.

REFERENCES

Paolo Albani & Berlinghiero Buonarroti. *Dictionnaire des langues imaginaires*. Paris, Les Belles Lettres, 2010, p. 576.

William Auld (ed.), *Esperanta antologio: Poemoj 1887-1981*. Rotterdam: Universala EsperantoAssocio, 1984, p. 887.

William Auld, *Kulturo kaj Internacia Lingvo*, Rotterdam, Universala Esperanto Associo, 1986, p. 150.

Henri Béjoint, « Informatique et lexicographie de corpus : les nouveaux dictionnaires », *Revue Française de Linguistique Appliquée*, Vol. 22/1, 2007, p. 7-23.

Benjamin K. Bergen, « Nativization processes in L1 Esperanto », *Journal of Child Language*, Vol. 28, 2001, p. 575-595

Detlev Blanke, *Internationale Plansprachen. Eine Einführung*. Berlin, Akademie-Verlag, 1985, p408.

Wera Blanke, « Terminological standardization - its roots and fruits in planned languages. », In Klaus Shubert & Dan Maxwell (eds.), *Interlinguistics: Aspects of the Science of Planned Languages,* Berlin, Mouton de Gruyter, 1989, p. 277-292.

Wera Blanke (ed.). *Esperanto - Terminologie und Terminologiearbeit. Esperanto - terminologio kaj terminologia laboro*. New York: Mondial, 2008, 156.

Osmo Buller, « Malkresko de la membraro kiel defio al UEA. » (Drop in membership: a challenge for the Universal Esperanto Association). *Esperantolehti -Gazeto de la Esperanto Asocio de Finnlando,* junio 2013,
Available at: http://www.liberafolio.org/2013/malkresko-de-la-membraro-kiel-defio-al-uea, consulted July 2013.

Oksana Viktorovna Burkina, *Proiznositel'naj a norma v Esperanto. Avtoreferat dissertacii na soiskanie uèenoj stepeni kandidata filologi èeskich nauk,* Sankt-Peterburg, Sankt-Peterburgskij gosudarstvennyi universitet, Special'nost' 10.2.19– Torija jazyka, 2009, p. 26.

Bernard Comrie, « Natural and Artificial International Languages: a Typologist's Assessment », *Journal of Universal Language,* Vol. 3, 1996, p. 35-55.

Renato Corsetti, « A Mother Tongue Spoken Mainly by Fathers. » *Language Problems and Language Planning,* Vol. 20 n° 3, 1996, p. 263-273.

Louis Couturat & Léopold Leau, *Lingvo internaciona di la delegitaro (Ido)*. Stuttgart; Frank,

1908, p. 24.
Natalia Dankova, *Temporalité en esperanto. Etude du transfert,* Thèse de doctorat, Dir. Clive Perdu, Université Paris 8, 1997, p.436.
Probal Daŝgupta, « Degree words in Esperanto and categories in Universal Grammar », In Klaus Shubert & Dan Maxwell (eds.), *Interlinguistics: Aspects of the Science of Planned Languages*, Berlin, Mouton de Gruyter, 1989, p. 231-248.
Probal Daŝgupta, « Idiomaticity and Esperanto texts: an empirical study. » *Linguistics,* Vol. 31/2, 1993, p. 367-386.
Michel Duc Goninaz, « La problemo de la normo en Esperanto.» In Ryszard Rokicki (red.). *Acta Interlinguistica. 12-a Scienca Interlingvistika Simpozio,* Varsovio, 27 aprilo – 01 majo, 1984, p. 57-65.
Michel Duc Goninaz, *La Nova Plena Ilustrita Vortaro de Esperanto*, Paris, Sennacieca Acocio Tutmonda, 2002, p. 1268.
John Edwards, *Multilingualism*, London, Routledge, 1994, p.272.
Mark Fettes, « The Esperanto community: a quasi-ethic linguistic minority? », *Language Problems & Language Planning,* Vol. 20, n° 1, 1996, p. 53-59.
Mark Fettes, « Interlinguistics and the Internet », *Language Problems & Language Planning,* Vol. 21, n° 2, 1997, p. 170-176.
Sabine Fiedler, « Standardization and self-regulation in an international speech community: the case of Esperanto. » *International Journal of the Sociology of Language,* Vol. 177, 2006, p. 67–90.
Sabine Fiedler, « Phraseology in Planned Languages. » In Harald Burger (ed.), *Phraseologie / Phraseology: ein internationales Handbuch zeitgenössischer Forschung / an international Handbook of Contemporary Research*, Vol. 2, 2007, p. 779-787
Peter Forster, *The Esperanto Movement.* Den Haag, Mouton, 1982, p. 419.
Robert C. Gardner & Wallace E. Lambert. *Attitudes and motivation in second language learnin*g, Rowley, MA: Newbury House, 1972, p. 313.
Christopher Gledhill, *The Grammar of Esperanto: A Corpus-Based Description,* München, Lincom Europa, 2000, p. 151.
Christopher Gledhill, « Les constructions verbo-nominales en français et en espéranto : un cas spécifique de glissement phraséologique. » *Zeitschrift für Französische Sprache und Literatur,* Vol. 36, 2008, 71-84
Christopher Gledhill, « Frazeologio, kunokazemo kaj leksiko-gramatiko. » (Phraséologie, collocation et lexicogrammaire). In Detlev Blanke & Ulrich Lins (eds.), *La arto labori kune. Festlibro por Humphrey Tonkin,* Rotterdam, Universala Esperanto-Asocio, pp. 323-338.
Michael A.K Halliday, *An Introduction to Functional Grammar*. London, Arnold, 1985, p. 387.
Michael A.K Halliday. « Language as system and language as instance: the corpus as a theoretical construct. » In Jan Svartvik (ed.) *Directions in corpus linguistics.* Stockholm. 1992, pp61-77.
Einar Haugen, « Dialect, language, nation ». In J. B. Pride & Janet Holmes (eds), *Sociolinguistics,* Harmondsworth: Penguin, 1972, pp.97–111.
Reinhard Haupenthal (ed.), *Eugen Wüster. Esperantologiaj Studoj: Memora Kolekto,* Antwerp, Stafeto & T. Kehler, 1978, p. 254.
Franz Josef Hausmann, & Peter Blumenthal. « Présentation: collocations, corpus, dictionnaires. » *Langue française,* Vol. 150, 2006, p. 3-13.
Wim Jansen, « The learnability of the reflexive in Esperanto », *Linguistics in Amsterdam,* Vol. 5, 2012, p. 57–80.

Wim Jansen, « Esperanto parts of speech in functional discourse grammar », *Linguistics*, Vol . 51, Issue 3, 2013, p. 611–652.

Kálmán Kalocsay & Gaston Waringhien, *Plena analiza gramatiko de Esperanto,* 5a eldono, Rotterdam, Universala Esperanto-Asocio, 1985, p. 599.

Aleksander Korĵenkov, *Historio de Esperanto.* (Serio *Scio no. 5*)*,* Kaliningrado, Editions Sezonoj, 2005, p. 128

Katalin Kováts (red.), *Komuna Eŭropa Referenckadro por lingvoj: lernado, instruado, pritaksado,* Rotterdam, Universala Esperanto-Asocio, 2007, p.262. (Available in English, French and Esperanto at: http://www.edukado.net/ekzamenoj/referenckadro).

Andrew Large, *The Artificial Language Movement,* Oxford, Blackwell, 1985, p.256

Wei Li, « Esperanto inflection and its interface in HPSG », *Working Papers of the Linguistics Circle* 13, 1996, p. 71-81.

Jouko Lindstedt, « Native Esperanto as a test case for natural language. » In Mickael Suominen, Antti Arppe, Anu Airola, Orvokki Heinämäki, Matti Miestamo, Urho Määttä, Jussi Niemi, Kari K. Pitkänen & Kaius Sinnemäki (eds.), *A Man of Measure: Festschrift in Honour of Fred Karlsson on his 60th Birthday,* Turku, Linguistic Association of Finland, 2006, p. 47–55.

Jouko Lindstedt, « Esperanto – an East European contact language? » In Christian Voß & Alicja Nagórko (eds.), *Die Europäizität der Slawia oder dies Slawizität Europas. Ein Beitrag der kultur- und sprachrelativistischen Linguistik* (Studies on Language and Culture in Central and Eastern Europe, 2), München / Berlin, Otto Sagner, 2009, p. 125-134.

Jonathan Pool & Bernard Gofman, « Linguistic artificiality and cognitive Competence. » In Klaus Schubert (ed.) *Interlinguistics Aspects of the Science of Planned Languages*, Trends in Linguistics 42, Berlin / New York, Mouton de Gruyter, 1989, p. 145-156.

Arika Okrent, *In the Land of Invented Languages,* New York, Spiegel & Grau, 2009, p352.

Mikael Parkvall, « How European is Esperanto? A typological study », *Language Problems & Language Planning,* Vol. 34, n° 1, 2010 , pp. 63-79.

Claude Piron, *Espéranto: l'image et la réalité*, Cours et études de linguistique contrastive et appliquée, n° 66, Paris, Université de Paris 8-Vincennes, 1986, p44.

Claude Piron, « A few notes on the evolution of Esperanto.» In Klaus Schubert (ed.) *Interlinguistics. Aspects of the Science of Planned Languages*, Trends in Linguistics 42, Berlin / New York, Mouton de Gruyter, 1989, p. 129-142.

Suzanne Romaine, « Revitalized languages as invented languages.» Michael Adams (ed.) *From Elvish to Klingon: Exploring Invented Languages,* Oxford, Oxford University Press, 2011, p. 185-225.

Dan Savatovsky, « Les linguistes et la langue internationale (1880-1920) ». *Histoire Epistémologie Langage, Vol.* 11/2, 1989, p 37-65.

Klaus Schubert, « An unplanned development in planned languages: A study of word grammar ». In Klaus Schubert (ed.) *Interlinguistics. Aspects of the Science of Planned Languages*, Trends in Linguistics 42, Berlin / New York, Mouton de Gruyter, 1989, p. 249–274.

Klaus Schubert, « Optimumigo de la internacia komunikado: interlingvistika rigardo al la nova industria lingvokonstruado. » In Sabine Fiedler & Haitao Liu (eds.) *Studien zur Interlinguistik, Festschrift für Detlev Blanke zum 60. Geburtstag,* 2001, p 172-184.

Bruce Arne Sherwood, « Statistical analysis of conversational Esperanto, with discussion of the accusative. » *Studies in the Linguistic Sciences,* Vol. 12, 1982, p. 165-182.

John McH Sinclair, *Corpus, Concordance, Collocation,* Oxford, Oxford University Press, 1991, p. 179.

Arden Smith, « Confounding Babel: International Auxiliary Languages. », Michael Adams (ed.) *From Elvish to Klingon: Exploring Invented Languages,* Oxford, Oxford University Press, 2011, p. 17-48.

Frank Stocker, *Wer spricht Esperanto? Kiu parolas Esperanton?* München, Lincom Europa, 1996, p.

Gordon Tucker, « Between lexis and grammar: towards a systemic functional approach to phraseology.», In Christian Matthiessen, Ruqaiya Hasan & Jonathan Webster (eds.) *Continuing Discourse on Language: A Functional Perspective,* Vol. 2. London: Equinox, 2007, p. 954-977.

Marc van Oostendorp, « Syllable structure in Esperanto as an instantiation of universal phonology », *Esperantologio / Esperanto Studies,* Vol. 1, 1999, p. 52-80.

Kees Versteegh, « Esperanto as a first language: Language acquisition with a restricted input.» *Linguistics,* Vol. 31/3, 1993, p. 539-555.

John Wells, *Lingvistikaj aspektoj de Esperanto*. (2a eldono), Rotterdam, Universala Esperanto Asocio, 1989, p. 76.

Bertilo Wennergren, *Plena Manlibro de Esperanta Gramatiko,* El Cerrito, Esperanto-League of North America, 2005, p. 696

John Wilkins, *An Essay towards a Real Character, and a Philosophical Language,* London, Gillibrand & Martin, 1668, p640.

Marina Yaguello, *Les langues imaginaires : mythes, utopies, fantasmes, chimères et fictions linguistiques*. Paris, Seuil, 2006, p.356.

Ludwik L. Zamenhof, *Международный языкъ. Предисловіе и полный учебникъ. (Meĵdunarodnij jazik. Predislovie i polnij uĉebnik, Internacia lingvo. Antaŭparolo kaj plena lernolibro),* Varšava, Tipo-Litografija X. Kel'tera, 1887, p.40. [In Huĝimoto et al. (eds.) *Unuaj libroj por esperantistoj*. Kyoto, Eldonejo Ludovikito, Vol. 1., 1973].

Ludwik L. Zamenhof, *Fundamenta Krestomatio,* Paris, Hachette, 1903, p. 463.

Ludwik L. Zamenhof, *Fundamento de Esperanto,* Paris, Hachette, 1905, p. 96.

Ludwik L. Zamenhof, *Lingvaj respondoj,* Paris, Esperantista Centra Librejo, 1936, p. 116.