Vol. 17.2 2019

Research in Language

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

RiL Vol. 17.2 2019

Research in Language

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

ŁÓDŹ 2019

EDITORS
Iwona Witczak-Plisiecka, Ewa Waniek-Klimczak, Jan Majer
University of Łódź, Faculty of Philology
Department of English Language and Applied Linguistics
90-236 Łódź, Pomorska 171/173

All papers published in "Research in Language" have been blind-reviewed
by at least two independent readers

INITIATING EDITOR
Agnieszka Kałowska

EDITING AND TYPESETTING
Michał Kornacki

COVER DESIGN
Katarzyna Turkowska

Printed directly from camera-ready materials provided to the Łódź University Press

Published by Łódź University Press
First Edition. W.09896.20.0.C
Printing sheets 6.875

ISSN 1731-7533
e-ISSN 2083-4616

Łódź University Press
90-131 Łódź, Lindleya 8
www.wydawnictwo.uni.lodz.pl
e-mail: ksiegarnia@uni.lodz.pl
phone (42) 665 58 63

CONTENTS

THE VOWEL SYSTEM OF PODHALE GORALIAN*

JERZY RUBACH
University of Iowa, USA, and University of Warsaw, Poland
jerzy-rubach@uiowa.edu; jerzy.rubach@uw.edu.pl

TOMASZ ŁUSZCZEK
University of Warsaw, Poland
t.luszczek@student.uw.edu.pl

Abstract
This paper is a report on the phonological research done in the past two years investigating Podhale Goralian. The data are drawn from our informants in Dzianisz.
The paper establishes the system of surface contrasts in Goralian and identifies instances of complementary distribution. It is claimed that the renowned Podhale Archaism is no longer represented by the vowel [i]. Rather, the vowel has retracted to the central vowel [ɨ]. The original [ɨ], on the other hand, has lowered and fronted and is now best regarded as tense [e]. These transitions of vowels pose challenges for a phonological analysis. A sample of such analysis (Final Tensing) is shown in the framework of Optimality Theory.

Keywords: Optimality Theory, Podhale Goralian, phonology, Polish dialects

1. Introduction

This paper is a report on the research done in the past two years investigating Podhale Goralian, a dialect of Polish spoken in the Podhale region of the Tatra Mountains in southern Poland. The data are drawn from our informants in Dzianisz, a village located about 10 km north-west of Zakopane. We take Dzianisz to be representative of the Podhale region and refer to the dialect as Podhale Goralian. Our fieldwork focused on the generation of speakers who are less than 50 years old. This is important because it may explain why our data differ from those reported in the literature in some significant ways.

The data come from an analysis of recorded interviews that included both free conversation (story telling) and reading of word lists. The recordings were transcribed independently by two professional phoneticians. Much data came

* We would like to thank the three reviewers of this paper for their comments and suggestions. We have a special debt of gratitude to our native speaker consultants for their help: Anna Łuszczek, Bogusława Łuszczek, Małgorzata Łuszczek, Wojciech Łuszczek, Józef Bąk, Andrzej Kamiński, Stanisław Kamiński, Anna Krupa, Józef Kula, Dorota Pająk, Katarzyna Pająk, Martyna Pająk, Władysław Pająk, Tomasz Płaza, Leszek Stasiwolak, Monika Szwajnos, and Andrzej Zarycki.

directly from one of the authors of this paper, Tomasz Łuszczek, who was born and raised in Dzianisz. He is not only a native speaker of the dialect but also a trained linguist and a phonetician.

Podhale Goralian is the most prominent and best known of the Goralian dialects. There is a huge literature about Podhale culture and language, but all of it is descriptive and not anchored in any theoretical framework. Reports on Podhale Goralian begin with Kopernicki (1875), Kosiński (1884), Kryński (1884), Stopka (1897, 1911), through Nitsch (1915), Małecki (1928) to Kąś (2015). The focus of this research is descriptive and history-driven, addressing, among other things, the question of what properties of Old Polish have been retained in Podhale Goralian.

This paper is a synchronic study. Section 2 introduces the system of vowels and shows basic contrasts. Section 3 looks at the current state of Podhale Archaism. Section 4 presents our proposal of the spelling system for Podhale Goralian. Section 5 is a sample analysis of Final ɔ-Tensing in terms of Optimality Theory. Section 6 concludes the paper with a summary of the most important results.

2. Vowel System

The vowels of Podhale Goralian are best illustrated against the background of the Standard Polish vowels. The cardinal vowels diagram in (Figure 1) is cited from Biedrzycki (1974: 28).[1]

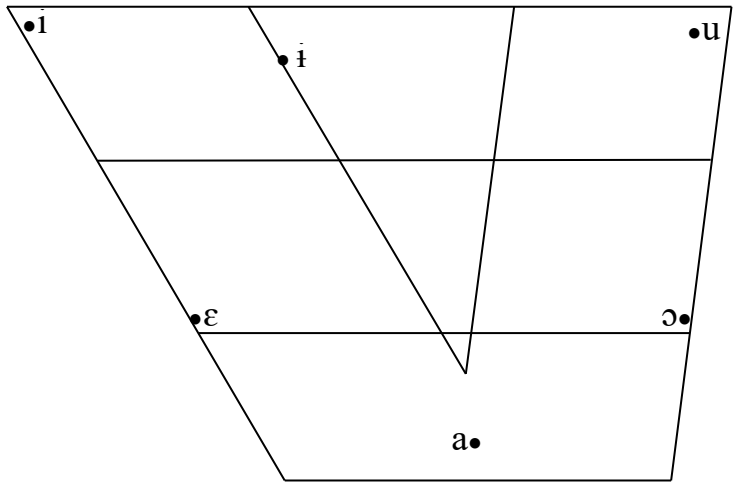


Figure 1. Vowels of Standard Polish

[1] We ignore nasal vowels and their decompositions in both Standard Polish and Podhale Goralian.

The vowels of our speakers from Dzianisz are different in both quantity and quality (for some vowels).

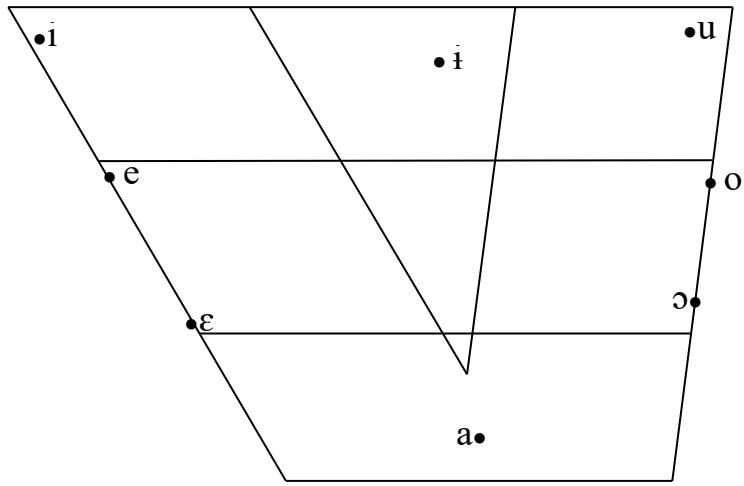


Figure 2. Vowels of Podhale Goralian

High vowels [i] and [u] are identical in (Figure 1) and (Figure 2). The central vowel [ɨ] is different in detail. A comparison of minimal pairs, such as *syn* [sɨn] 'son' (Podhale Goralian) and *syn* [sɨn] 'son' (Standard Polish), shows that the Podhale Goralian [ɨ] is more retracted than the Standard Polish [ɨ] and sounds like the Russian [ɨ] in *syn* 'son'.

Podhale Goralian diverges from Standard Polish in that it has four rather than two mid vowels.

(3)	*nom.sg.*	*gen.sg.*	*gloss*
	chleb [xlep]	chleb+a [xlɛba]	'bread'
	śnieg [ɕɲek]	śnieg+a [ɕɲɛga]	'snow'
	próg [prok]	prog+a [prɔga]	'threshold'
	Bóg [bok]	Bog+a [bɔga]	'God'

The additional mid vowels, tense [e] and tense [o], are typical for regional dialects of Polish, for example, they are attested in Kurpian (Rubach 2011). They existed in the standard dialect in Middle Polish, but disappeared in the 19th c.

Tense [e] and [o] have a firm status as contrastive segments. They can be regarded as derivable by rule (Final Syllable Tensing) in the examples cited earlier in (3). This is not true for many other words because tense [e o] do not exhibit alternations.

(4)	*nom.sg.*	*nom.pl.*	*gloss*
	kołnierz [e]	kołnierz+e [e][2]	'collar'
	lek [e]	lek+i [e]	'drug'
	but [o][3]	but+y [o]	'shoe'
	król [o]	król+e [o]	'king'

Also: *apteka* [apteka] 'drug store', *chlew* [xlef] 'pigsty', *góra* [gora] 'mountain'

In these examples [e o] occur both finally and medially, so they cannot be derived by Final Syllable Tensing. We conclude that //e o// are underlying segments.

The third historically tense vowel is *a*. It occurs in both alternating (5a) and non-alternating (5b) contexts.

(5)		*gen.sg.*	*nom.sg.*	*gloss*
	a.	przykład+u	przykłod	'example'
		dziad+a	dziod	'idler'
		pan+a	pon	'tourist'
		Jan+a	Jon	'John'
	b.	ptok+a	ptok	'bird'
		stow+u	stow	'pond'
		plebon+a	plebon	'priest'
		dzbon+a	dzbon	'pitcher'

As (5) documents, the Podhale Goralian reflex of the historical tense [ɑ] is *o*. Descriptive sources (for example, Kąś 2015) characterize this *o* as an intermediate vowel between [ɑ] and [ɔ]. This characterization is not true for our informants from Dzianisz.

(6)	*Standard Polish*	*Podhale Goralian*	*gloss*
	a. Polak	Polok [pɔlɔk]	'Pole'
	b. kowal	kowol [kɔwɔl]	'blacksmith'

Our informants claim that they perceive the vowels in (6) as exactly the same, and we concur with this judgement. Linguistically speaking, the result is that the neutralization process of the historical [ɑ] with the historical [ɔ] has been completed. As might be expected, the homophony leads to phonological problems because the [ɔ] from [ɑ] sometimes alternates with [a], as in (5a), and sometimes does not, as in (5b). Addressing this issue is matter of future research.

---

[2] We argue in Section 4 that examples such as this one should be written *kołńyrz*.

[3] In Podhale Goralian, the Standard Polish word *but* 'shoe' is pronounced with [o], not with [u]. Consequently, as we explain in Section 4, it is spelled *bót*.

Finally, let us note that the loss of distinction between the two kinds of *o* shown in *Polok* (6a) is a recent change. Kąś (2015) asserts that the vowels are different, though likely to be non-distinct with speakers of the younger generation.

To conclude, the vocalic system of Podhale Goralian is richer than that of Standard Polish, but both are rather typical as phonological systems.

(7) a. Standard Polish

i	ɨ	u
ɛ		ɔ
	a	

b. Podhale Goralian

i	ɨ	u
e		o
ɛ		ɔ
	a	

A featural classification of the system in (7b) poses certain problems. First, with three clearly front vowels [i e ɛ] and three evidently back vowels [u o ɔ], the question is about the status of the two central vowels [ɨ a] *vis-á-vis* the feature [±back]. Second, it needs to be determined how to accommodate the additional contrast that obtains in the region of mid vowels.

With regard to the first question, the phonetic facts come to no avail because, technically speaking, [ɨ] and [a] are central. However, the answer is readily provided by the phonological behavior of these vowels.

(8) mróz [mros] ‘frost’ (nom.sg.)
a. [mrɔʑ+ik] (dim.): z→ʑ /_ i
b. [mrɔʑ+ɛ] (loc./voc.sg.): z→ʑ /_ ɛ[4]
c. [mrɔz+ɨ] (nom.pl.), [mrɔz+a] (gen.sg.)[5]

It is common knowledge in phonology that palatalization is caused by front, that is, [-back] vowels. The Goralian data in (8a–b) corroborate this claim: the vowels [i] and [ɛ] trigger the change *z→ʑ*, which, in terms of distinctive features, involves a change in the specification of the feature [±back]: [+back] → [-back]. In contrast to the forms containing [i] and [ɛ] in (8a–b), the forms in (8c) containing [ɨ] and [a] do not cause palatalization: we see [z] rather than [ʑ] in [mrɔz+ɨ] (nom.pl.) and [mrɔz+a] (gen.sg.). Given this phonological evidence, we conclude that the vowels [ɨ] and [a] must be classified as [+back].

The tense – lax vocalic distinction in the class of mid vowels is readily captured by [±advanced tongue root] ([±ATR], henceforth). In the class of mid vowels, [+ATR] designates upper mid vowels while [-ATR] refers to lower mid vowels.

Pulling together the observations made about Goralian vowels, we arrive at the following feature chart.

[4] This vowel surfaces as tense [e] in fact. It remains unclear whether the change is due to an independently motivated phonological rule or, rather, an allophonic process raising vowels in the vicinity of palatalized consonants.

[5] The gen.sg. ending in Goralian *mróz* ‘frost’ is *–a*, not *–u*.

(9) Podhale Goralian vowels[6]

	i	ɨ	u	e	o	ɛ	ɔ	a
high	+	+	+	–	–	–	–	–
low	–	–	–	–	–	–	–	+
back	–	+	+	–	+	–	+	+
ATR	+	+	+	+	+	–	–	–
round	–	–	+	–	+	–	+	–

The vowels in (7b) are all underlying segments because they are contrastive. Some contrasts are obvious, involving many different properties; other contrasts are more subtle, as the following examples demonstrate.

(10) *e* – *ɛ*, [lek] ‘drug’ – [stɛk] ‘steak’
[vʲes] ‘you know’ – [vʲɛɛ] ‘village’[7]
[pɔɕtɕel] ‘bed’ – [tɕɛlɛ] ‘calf’
e – *ɨ*, [sen] ‘dream’ – [sɨn] ‘son’
[gʒep] ‘bury’ (imper.) – [gʒɨp] ‘mushroom’
[ʃʧets] ‘guard’ – [ʃʧɨts] ‘cut’
i – *ɨ*, [nɔɕi] ‘he carries’ – [nɔsɨ] ‘noses’
[vɔʑi] ‘he transport’ – [vɔzɨ] ‘carts’
[kɔʑɨtsa] ‘mountain goal’ – [zɨtɕe] ‘life’
o – *u*, [bot] ‘shoe’ – [bus] ‘minibus’
[prok] ‘threshold’ – [truxwo] ‘corpse’
[dom] ‘home’ – [duk] ‘ghost’
o – *ɔ* [mots] ‘be able’ – [mɔts] ‘power’
[droga] ‘road’ – [drɔn] ‘drone’
[krol] ‘king’ – [krɔk] ‘step’

The phonological system in (7b) is extended further by diphthongs that originate either from underlying lax //ɔ// or from underlying tense //o// and derived tense /o/. The diphthongs are optional or rather phonostylistic.

(11) ɔ → uɔ [kɔt] or [kuɔt] ‘cat’
[kɔvɔl] or [kuɔvɔl] ‘blacksmith’
[pɔt] or [puɔt] ‘sweat’
[pɔ] or [puɔ] ‘after’

[6] Goralian contrasts lax and tense vowels only in the mid region, thus the vowels [i ɨ u a] are specified for [±ATR] in a redundant manner.

[7] The [ɛ] is raised by an allophonic or phonetic implementation process operating in the context of palatalized consonants, but the raising is not significant enough to obliterate the contrast with tense [e].

	[bɔtsɨtɕ] or [buɔtsɨtɕ] ‘hold a grudge’
	[nɔve] or [nuɔve] ‘new’
o → uo	[gora] or [guora] ‘mountain’
	[mleko] or [mlekuo] ‘milk’
	[koɲ] or [kuoɲ] ‘horse’
	[vor] or [vuor] ‘sack’ (augment.)
	[bo] or [buo] ‘why’
	[dom] or [duom] ‘home’

The point of interest is that the diphthongs have the status of short rather than long vowels, which is natural as Podhale Goralian is not a weight-sensitive dialect. The [u] onglide of the diphthong is subsegmental because native speakers perceive a difference between [wɔ] and [uɔ] as well as between [wo] and [uo].[8]

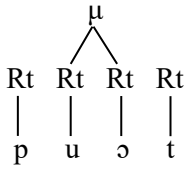
(12) a. *płot* [pwɔt] ‘fence’ is not the same as *pot* [puɔt] ‘sweat’
kłos [kwɔs] ‘ear of corn’ is not the same as *kos* [kuɔs] ‘blackbird’
kładł [kwɔt] ‘he laid’ is not the same as *kot* [kuɔt] ‘cat

b. *chłód* [xwot] ‘cold’ is not the same as *chód* [xuot] ‘gait’
tło [two] ‘background’ is not the same as to [tuo] ‘this’
kłonica [*kwoɲ*itsa] ‘rod’ is not the same as *konnica* [*kuoɲ*ɲitsa] ‘cavalry’

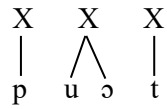
The [uɔ] and [uo] strings are diphthongs, not sequences of [u] and [ɔ] or [o] because the words in (12) are perceived by native speakers as monosyllabic.[9] Thus, [puɔt] ‘sweat’ (nom.sg.) is one syllable, not two, while [puɔt+u] (gen.sg.) is two syllables, not three syllables.

We conclude that Podhale Goralian has short diphthongs. Technically, they are represented as two root nodes linked to a single mora or X-slot, depending on the skeletal theory. Rt stands for the Root node.[10]

(13) a. moraic theory



b. X-skeletal theory



The occurrence of short diphthongs is interesting typologically because short diphthongs are extremely rare. The only well-known case is Spanish, for which

[8] This is based on the judgements made by native speakers, who were asked whether pairs such as *płot* ‘fence’ and *pot* ‘sweat’ in (12a) sounded the same or different. The judgement was that they were different. The same conclusion was reached independently by the transcribers.

[9] [kwoɲitsa] and [kuoɲɲitsa] have three syllables each.

[10] We assume the Halle-Sagey feature geometry (Halle 1992, Sagey 1986).

they have been reported by Clements and Keyser (1983). While the occurrence of short diphthongs in Podhale Goralian is a firm fact,[11] the circumstances under which they are actually used are not entirely clear. The general observation is that diphthongs are used for emphasis in an obligatory way. Other stylistic use of diphthongs is not clear and requires research.

3. Podhale Archaism

Goralian is renowned in Polish linguistic literature for the so-called Podhale Archaism (*archaizm podhalański*). The matter has drawn the attention of linguists ever since the first studies of Podhale Goralian in the 19th c., from Kosiński (1884), Nitsch (1915), Małecki (1928) to Dejna (1973) and Decaux (1973).

Historical grammarians, for example, Rozwadowski (1915), Stieber (1952, 1973) are in agreement that until the 16th c., Polish stridents were soft, that is [-back], segments: [tsʲ dzʲ ʧʲ dʒʲ ʃʲ ʒʲ].[12] Consequently, they combined with the front vowel [i] rather than with the back vowel [ɨ],[13]as the following examples show.

(14)		*15th c. Polish*	*modern Polish*	*gloss*
	chłopcy	[xɫɔptsʲi]	[xwɔptsɨ]	‘boys’
	drodzy	[drɔdzʲi]	[drɔdzɨ]	‘dear’ (nom.pl.)
	czy	[ʧʲi]	[ʧɨ]	‘if’
	szyć	[ʃʲitɕ]	[ʃɨtɕ]	‘sew’
	żyto	[ʒʲitɔ]	[ʒɨtɔ]	‘rye’

The original soft [tsʲ dzʲ ʧʲ dʒʲ ʃʲ ʒʲ] hardened, that is, became [+back], in the 16th c. The fallout of Hardening is different in different Slavic languages and dialects. For example, Ukrainian has retained soft [tsʲ] (Bilodid 1969) while Russian has kept soft [ʧʲ] (Avanesov 1968). Lower Sorbian is still different: it has retained soft posteriors [ʧʲ dʒʲ ʃʲ ʒʲ] (Schuster-Šewc 1968). Standard Polish hardened all stridents, the consequence being that [i] was replaced by [ɨ].[14]

(15) Hardening tsʲ dzʲ ʧʲ dʒʲ ʃʲ ʒʲ → ts dz ʧ dʒ ʃ ʒ[15]

Hardening did not affect Podhale Goralian, so examples such as those in (12) have retained their original [i]. It is this fact that is called Podhale Archaism (*archaizm*

[11] The literature talks about ‘labialized *o*’ that can occur word-medially; see, for example, Kąś (2015).

[12] [dʒʲ] spirantized to [ʒʲ].

[13] Central vowels are characterized as [+back] in generative phonology.

[14] For discussion, see Rubach (2003, 2017).

[15] The generalizations in this section are descriptive. A sample of a formal analysis is provided in Section 5.

podhalański): the retention of [i] is archaic because it reflects the state of Polish in the 15th c. The actual pronunciation of the examples in (12) diverges from the 15th c. pronunciation in one way: Podhale Goralian, like many other regional dialects of Polish (for example, Kurpian), has developed a rule that we will call Dentalization (*mazurzenie*).

(16) Dentalization (*mazurzenie*) ʧ ʤ ʃ ʒ → ʦ ʣ s z[16] context-freely

The combined effect of Dentalization and the retention of archaic [i] yields the following pronunciation of the examples in (14).

(17) Podhale Goralian Archaism

a. Archaism:
chłopcy [xwɔptsʲi] 'boys'
drodzy [drɔdʑi] 'dear' (nom.pl.)

b. Archaism plus Dentalization:
czy [tsʲi] 'if'
szyć [sʲitɕ] 'sew'
żyto [zʲitɔ] 'rye'[17]

Małecki (1928) notes that the [i] of Podhale Archaism is not the same as the [i] in other contexts, such as in *pić* 'drink'. He claims that it is an intermediate vowel between [i] and [ɨ], so probably [ɪ], which he proposes to represent in the spelling with a new letter: *ý*. Małecki (1928) makes two other important points. First, he notes that the territory on which Podhale Archaism operates starts shrinking. Second, the villages that exhibit the Archaism, Dzianisz including, exhibit it in a robust way: the Archaism is very much alive and productive.

Our fieldwork carried out in Dzianisz almost a hundred years later shows that Podhale Goralian has undergone a dramatic change since the time when Małecki collected his data: the [i] or [ɪ] has completely disappeared in the language of Goralians who are 50 years old or younger. What we find is [ɨ] in the very same Podhale Archaism words.

(18) Change of Podhale Archaism in Dzianisz

chłopcy 'boys' [xwɔptsʲi] → [xwɔptsɨ]

[16] Less frequently, ʧ ʤ ʃ ʒ → tɕ ʣ ɕ ʑ.

[17] It is probably Dentalization that is responsible for the extension of the Archaism to words with the original [s z] because the original [s z] and the [s z] from Dentalization are non-distinct. The consequence is that Podhale Archaism now encompasses also words such as *syn* 'son' and *zysk* 'profit'.

drodzy ‘dear’ (nom.pl.)	[drɔdzʲi] → [drɔdzɨ]
czy ‘if ’	[tsʲi] → [tsɨ]
szyć ‘sew’	[sʲitɕ] → [sɨtɕ]
żyto ‘rye’	[zʲitɔ] → [zɨtɔ][18]

The ‘new [ɨ]’ in (18), shown earlier in diagram (1b), is definitely not a front but a central vowel and sounds much like Russian [ɨ]. The ‘new [ɨ]’, exactly like the original Podhale Archaism, extends to the context of historical [s z], such as in *syn* [sɨn] ‘son’ and *zysk* [zɨsk] ‘profit’.

The rise of ‘new [ɨ]’ leads to the question of what has happened to the ‘old [ɨ]’ in contexts other than those defined by Podhale Archaism. The answer is that the ‘old [ɨ]’ has totally disappeared because it has collapsed with tense [e].

(19) Vowel Transition [ɨ] → [e]

Since Vowel Transition did not occur in the Podhale Archaism contexts, it can be surmised that Podhale Archaism blocked or undid the effects of Vowel Transition. A theoretical analysis of this issue awaits future research.

As already noted, Podhale Goralian has tense [e] in words inherited with tense [e] from Old Polish, such as *chlew* [xlef] ‘pigsty’, *grzech* [gʒex] ‘sin’ and *sierp* [ɕerp] ‘sickle’. The point of interest is that native speakers judge the following words as containing the same vowel in both syllables. Consequently, the transcription must be as given below.

(20) chlew+y [xleve] ‘pigsties’
grzech+y [gʒexe] ‘sins’
sierp+y [ɕerpe] ‘sickles’

At the same time the vowels in the following words are judged as different, a fact that we note in the transcription.

(21) syty [sɨte] ‘full’
syny [sɨne] ‘sons’
czysty [tsɨste] ‘clean’

The facts in (21) fit into our story: after the Podhale Archaism consonants, here [s ts], we have [ɨ] while [e] for ‘old [ɨ]’ occurs elsewhere.

We wish to make a further observation. The speakers in Dzianisz have extended the classic Podhale Archaism context not only to the original [s z], exemplified by *syn* [sɨn] ‘son’ and *zysk* [zɨsk] ‘call’, but also to the posterior fricatives [ʃ] and [ʒ] that are the effects of *r*-Palatalization.

[18] The actual surface form is [zɨto], with the word-final //ɔ// changed to [o]. See Section 5.

(22) trzy ‘three’: [tʃɨ], not *[tʃe] or *[tʃʲi]
krzywy ‘crooked’ [kʃɨve], not *[kʃeve] or *[kʃʲive]
Rzym ‘Rome’ [ʒɨm], not *[ʒem] or *[ʒʲim]
przyroda ‘nature’ [pʃɨroda], not *[pʃeroda] or *[pʃʲiroda]

In sum, the descriptive generalization is that ‘new [ɨ]’ occurs in the following contexts.

(23) Distribution of ‘new [ɨ]’

> The vowel [ɨ] occurs after hard stridents [s z ʦ ʣ ʃ ʒ], regardless of whether the source is Dentalization, *r*-Palatalization or the underlying representation.

To conclude, somewhat paradoxically, the dramatic sound change *i* → *ɨ* has not eliminated Podhale Archaism, but made it different: the hallmark of Podhale Archaism is now the occurrence of [ɨ], not of [i]. The transition of ‘old [ɨ]’ to [e] poses huge challenges for phonological theory. For one thing, we have a front vowel (tense [e]) that does not cause Palatalization. These challenges constitute a research agenda for the future.

4. Spelling

We have now identified the relevant vowel types of Podhale Goralian. This leads to an important practical question: how should these distinctions be noted in the spelling?

There has been much debate about the spelling of Podhale Goralian. Stopka (1897) uses accented *ý* to spell the vowel of Podhale Archaism.[19] In a later publication (Stopka 1911), he rejects *ý* in favour of *ī*.

(24) Spelling of Podhale Archaism

Standard Polish	*Stopka (1897)*	*Stopka (1911)*	*gloss*
czysty	cýsty	cīsty	‘clean’
szyć	sýć	sīć	‘sew’
żyć	zýć	zīć	‘live’
syn	sýn	sīn	‘son’
zysk	zýsk	zīsk	‘profit’

[19] Kopernicki (1875) also uses *ý* in the same context to represent, as he puts it, *long y*.

Years later in his monograph, Małecki (1928) returns to *ý*, but assigns it a new quality: *ý* denotes an intermediate vowel between *i* and *y*. The letter *ý* is used still today and denotes an [i] type vowel (Kąś 2015), not an [ɨ] vowel. Given this tradition, we cannot retain *ý* to mean the Podhale Archaism vowel because we have determined that the vowel is [ɨ], not a front vowel of an [i] type. Our proposal is to introduce the letter *ȳ* to represent the central vowel [ɨ].

We go along with the widely spread tradition of writing *y* to mean tense [e] for the historical *chléb* [xlep] 'bread' and *chlyw* [xlef] 'sty', as in Dejna (1973). However, as explained in Section 3, we take all *y*'s to stand for [e], not only those *y*'s that correspond to the historical tense *é*. Dejna (1973) has a different view. He transcribes *brzég* 'shore', *rzéka* 'river' and *śniég* 'snow' as [bžyk], [žyka] and [śńyk], respectively, and claims that tense [e] has narrowed to [y].[20] This is the opposite of what we are saying. In our view, it is the [y] (our 'old [ɨ]) that has lowered to tense [e].

Consequently, we have the following spelling of Standard Polish words in Podhale Goralian. The words in (25) exemplify [e] and the Podhale Archaism [ɨ] discussed earlier.

(25)	*Standard Polish*	*Podhale Goralian*	*gloss*
a.	grzech [gʒɛx]	grzych [gʒex]	'sin'
	jeż [jɛʃ]	jyż [jeʃ]	'hedgehog'
	zlew [zlɛf]	zlyw [zlef]	'sink'
b.	czysty [ʧɨstɨ]	cȳsty [tsɨste]	'clean'
	szyć [ʃɨtɕ]	sȳć [sɨtɕ]	'sew'
	żyć [ʒɨtɕ]	zȳć [zɨtɕ]	'live'
	syn [sɨn]	sȳn [sɨn]	'son'
	zysk [zɨsk]	zȳsk [zɨsk]	'profit'

We agree with the established tradition across Polish regional dialects that tense [o] is simply written as *ó*. The representation of the diphthongs is a little problematic. It is common practice to use the letter *ô* to represent both [uɔ] as in *kot* [kuɔt] 'cat', written *kôt*, and word-initial epenthetic [w], as in *owca* [wɔftsa] 'sheep', written *ôwca*. The spelling of both postconsonantal [uɔ] and the onset-*o* sequence [wɔ] in the same way as *ô* is a problem because we believe that they are different phonological beings: [uɔ] is a short diphthong while [wɔ] is an epenthetic effect of the constraint ONSET that penalizes onsetless syllables. The distinction is not only theory-driven. As shown in Section 3, native speakers perceive [uɔ] and [wɔ] as different, so they need to be spelled differently. Our suggestion is to use *ŏ* (hacek *o*) for the lax diphthong [uɔ] and *ł* for epenthetic [w].[21] The tense

[20] [y] is the Slavic way of representing IPA [ɨ].

[21] We have decided against using *ô* because of the confusing content that authors assign to it. In addition to [uɔ] and [wɔ] just discussed, it has been suggested that *ô* should stand for plain [ɔ] (Kąś 2012).

diphthong [uo] derived from tense [o], as in [guora] ‘mountain’, has never been discussed in the literature before, so the spelling has never been suggested either. We propose to use the letter *ő* for [uo], so [guora] is written *gőra*.

The table in (26) summarizes the orthoepic relations between vowel letters and vowel sounds.

(26) Vowel letter – sound correspondences in Podhale Goralian

Letter	IPA Symbol	Description	Example
ȳ	[ɨ]	unrounded central vowel	sȳć [sɨtɕ] ‘sew’
y	[e]	mid tense unrounded front vowel	chlywy [xleve] ‘pigsties’
ó	[o]	Mid tense rounded back vowel	król [krol] ‘king’
ŏ	[uɔ]	short lax diphthong	kŏt [kuɔt] ‘cat’
ő	[uo]	short tense diphthong	gőra [guora] ‘mountain’

The other vowel letters, *i, u, e, o,* and *a*, represent the same sounds as in Standard Polish: [i u ɛ ɔ] and [a], respectively.

In addition to defining the representation of vowels, we propose to simplify the representation of prepalatals [ɕ ʑ tɕ dʑ ɲ]. To see the point, we need to first consider the spelling system for prepalatals in Standard Polish. Every prepalatal is spelled in three different ways.

(27) Standard Polish

a. Before a consonant or a word boundary, prepalatals are spelled with an accent over the letter.
   [ɕ]: kość ‘bone’, struś ‘ostrich’
   [ʑ]: groźba ‘threat’, paź ‘page’[22]
   [tɕ]: ćwiczenie ‘exercise’, kapeć ‘slipper’
   [dʑ]: dźwig ‘crane’, sledź ‘herring’
   [ɲ]: kończyć ‘end’, leń ‘lazy person’

b. Softness is marked not by an accent but by the letter *i* if *i* is followed by another vowel letter; in that case, *i* functions as a diacritic and is not pronounced, unless the preceding consonant is a labial.[23]
   [ɕ]: siano [ɕanɔ] ‘hay’
   [ʑ]: ziarno [ʑarnɔ] ‘grain’
   [tɕ]: ciasto [tɕastɔ] ‘cake’
   [dʑ]: działo [dʑawɔ] ‘cannon’
   [ɲ]: nie [ɲɛ] ‘no’

[22] Polish has Final Devoicing, so *paź* and *śledź* (below) are pronounced with a voiceless obstruent.

[23] A reviewer asks about the context of velars, as in *kiedy* ‘when’, *giez* ‘hadfly’ and *hiena* ‘hyena’. The answer is that *i* is pronounced as [j] after velar fricatives, but not after velar stops: [kʲɛdɨ], [gʲɛs], [xʲjɛna].

Labials: piasek [pʲjasɛk] ‘sand’, biały [bʲjawɨ] ‘white’, miasto [mʲjastɔ] ‘town’, wiadro [vʲjadrɔ] ‘bucket, parafia [parafʲja] ‘parish’

c. Prepalatals are written as plain consonants before *i* that is not followed by a vowel; so the *i* has two functions: (i) it is a nucleus and (ii) it is a diacritic denoting softness.

siwy [ɕivɨ] ‘gray’
zimny [ʑimnɨ] ‘cold’
cichy [tɕixɨ] ‘quiet’
dziki [dʑikʲi] ‘wild’
nic [ɲitsˌ] ‘nothing’

As noted, this is a complex system, with three ways of spelling the same sound. Particularly disturbing is the situation when the same sound in the same morpheme is spelled in two or even three ways.

(28) struś ‘ostrich’ (nom.sg.): [ɕ] spelled with an accent as *ś*, [struɕ]
strus+ia (gen.sg.)[24]: [ɕ] spelled with mute *i*, [struɕa]
niedźwiedź ‘bear’ (nom.sg.): final [dʑ][25] spelled *dź*, [ɲɛdʑvʲjɛdʑ]
niedźwiedz+ia (gen.sg.): the same [dʑ] spelled with mute *i*, [ɲɛdʑvʲjɛdʑa]
niedźwiedz+ic+a ‘she bear’: here [dʑ] is spelled *dzi*, where *i* is pronounced, [ɲɛdʑvʲjɛdʑitsa]

The paradox is that this complex system turns out to be not complex enough in order to represent different sounds in an adequate way. For example, the *si* in *sinus* ‘sine’, the *zi* in *Zanzibar* ‘Zanzibar’ and the *ci* in *na cito* ‘urgently’ falsely suggest that the pronunciation is *[ɕi], *[ʑi] and *[tɕi], respectively. The attested pronunciations are [sʲi], [zʲi] and [tsʲi].

All of these complexities and inadequacies are avoided if we adopt the following spelling convention.

(29) Spelling of prepalatals

The only way to spell [ɕ ʑ tɕ dʑ ɲ] is by using an accent, so by using the letters *ś, ź, ć, dź*, and *ń*, respectively.

Given this principle, the Standard Polish examples in (27b, c) and in (28) would be spelled as follows.

[24] The masculine forms *strus+i* and *niedźwiedz+i* (below) belong to the category illustrated by *niedźwiedz+ic+a* ‘she bear’ below, where *i* appears in two functions: it is pronounced and denotes palatalization

[25] We ignore Final Devoicing.

(30)

[ɕ]:	śanó [ɕano] ‘hay’	śiwy [ɕive] ‘gray’
[ʑ]:	źarnó [ʑarno] ‘grain’	źimny [ʑimne] ‘cold’
[tɕ]:	ćastó [tɕasto] ‘cake’	ćichy [tɕixe] ‘quiet’
[dʑ]:	dźałó [dʑawo] ‘cannon’	dźiki [dʑikʲi] ‘wild’
[ɲ]:	ńy [ɲe] ‘no’	ńic [ɲits] ‘nothing’
struś ‘ostrich’ (nom.sg.)		struś+a [struɕa] (gen.sg.)
ńedźwjydź ‘bear’ (nom.sg.)		ńedźwjydź+a [ɲɛdʑvʲjedʑa] (gen.sg.)
		ńedźwjydź+ic+a [ɲɛdʑvʲjedʑitsa] ‘she bear’

The troublesome examples, *sinus* ‘sine’, *Zanzibar* ‘Zanzibar’ and *na cito* ‘urgently’ are not a problem any longer: they are not pronounced with *[ɕi], *[ʑi] and *[tɕi] because they are not written **śinus, *Zanźibar* and **na ćito*. To conclude, we adopt the principle in (29), so the only correct way of spelling prepalatals in Podhale Goralian is to use accented letters, as shown in (30).

The principle adopted here that spelling should reflect the facts of pronunciation as closely as possible,[26] has a consequence for writing glides. The rule is simple: glides are written where they are heard. The glide [w] is written *ł* and the glide [j] is spelled *j*, as the following examples illustrate.

(31) Spelling of glides

a.

łosa [wɔsa] ‘wasp’	
łokó [wɔko] ‘eye’	
łowca [wɔftsa] ‘sheep’	
łokap [wɔkap] ‘cooker hood’	
łozwa [wɔzva] ‘echo’	
pałza [pawza] ‘pause’	
Ełrópa [ɛwropa] ‘Europe’	

b.

pjosek [pʲjɔsɛk] ‘sand’	
bjoły [bʲjɔwe] ‘white’	
mjastó [mʲjasto] ‘city’	
fjołek [fʲjɔwɛk] ‘violet’	
wjeś [vʲjɛś] ‘village’	
Ałstralja [awstralʲja] ‘Australia’	
intujicja [intujitsʲja] ‘intuition’	

The glide [w] can be written *ł* with impunity today because the classic pronunciation of *ł* as dark *l* [ɫ] hardly exists: it can be found only with the oldest speakers (in their 80’s and 90’s).

The final question concerns the correspondences in the spelling and pronunciation between the vowels of Standard Polish and the vowels of Podhale Goralian. The correspondences are represented as follows.

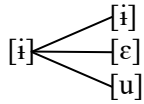
(32) Vowel correspondences: Standard Polish – Podhale Goralian

a. [i] → [i]

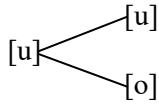
[26] Spelling should not be confused with transcription. There are certain details of pronunciation that need not, or even should not, be reflected in the spelling because then the spelling is not practical.

[e]

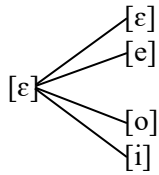
Standard Polish	*Podhale Goralian*	*gloss*
nos*i*ć	noś[i]ć	‘carry’
szuś*ci*ć	suść[e]ć	‘rustle’

b. 

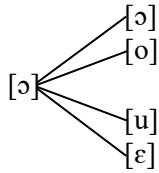
Standard Polish	*Podhale Goralian*	*gloss*
grz*y*b	grz[ɨ]b	‘mushroom’
t*y*le	t[ɛ]ló	‘this much’
wyr*y*wać	wyr[u]wać	‘pluck out’

c. 

Standard Polish	*Podhale Goralian*	*gloss*
p*u*kać	p[u]kać	‘knock’
dzi*u*ra	dź[o]ra	‘hole’

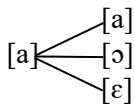
d. 

Standard Polish	*Podhale Goralian*	*gloss*
z*e*ro	z[ɛ]ró	‘zero’
chl*e*w	chl[e]w	‘sty’
tyl*e*	tel[o]	‘this much’
si*e*kiera	ś[i]kjyra	‘axe’

e. 

Standard Polish	*Podhale Goralian*	*gloss*
m*ł*ot	m[ɔ]t	‘hammer’

wiadr*o*	wjadr[o]	‘bucket’
k*o*mora	k[u]móra	‘chamber’
mi*o*tła	mj[ɛ]tła	‘broom’

f. 

Standard Polish	*PodhaleGoralian*	*gloss*
m*a*m*a*	m[a]m[a]	‘mother’
szp*a*k	śp[ɔ]k	‘starling’
tr*a*fić	tr[ɛ]fić	‘hit the mark’

5. Sample analysis

As already noted, working out the phonological system of Podhale Goralian is a huge and complex project. We exemplify the issue by looking at a single problem that appears to be straightforward: Final *o*-Tensing. The data come 100% from our fieldwork; the generalizations that emerge from the data are new and have never been reported in the literature.

(33) Final tense [o]

neuter nom.sg.: lat+*ó* [o] ‘summer’, błot+*ó* [o] ‘mud’, lustr+*ó* [o] ‘mirror’[27]
fem. voc. sg.: bab+*ó* [o] ‘woman’, idjotk+*ó*[28] [o] ‘idiot’, mam+*ó* [o] ‘mother’
adverb. suffix: hrub+*ó* [o] ‘thick’ (Adv.), tłust+*ó* [o]‘fat’ (Adv.), mał+*ó* [o] ‘little’
Adj. masc. gen.sg.: hrub+eg*ó* [o] ‘thick’ (Adv.), tłust+eg*ó* [o]‘fat’ (Adv.), mał+eg*ó* [o] ‘small’ (Adv.)

Also: b*ó* [o] ‘because’, h*ó* [o] (exclamation word used with horses)

It appears that what we are looking at is a rule of Final ɔ-Tensing. Schematically:

(34) Final ɔ-Tensing: ɔ → o / — #

[27] Since generative phonology reasons in term of morphemes rather than words, these three nouns represent one example: the nom.sg. ending.

[28] The spelling with *j* rather than with *i* follows from the phonetic principle introduced in the preceding section that glides are written where they are heard.

Final *ɔ*-Tensing is challenged by the following examples:

(35) Final lax [ɔ]

verbalizing suffix:	cȳt+*o* [ɔ] 'he reads', śpjyw+*o* [ɔ] 'he sings', gr+*o* [ɔ] 'he plays'
Adj. fem. nom.sg.:	hrub+*o* [ɔ] 'fat', tłust+*o* [ɔ] 'fat', wysók+*o* [ɔ] 'tall'
fem. nom.sg.:	plebańj+*o*[29] [ɔ] 'presbytery', litańj+*o* [ɔ] 'litany', łoperacj+*o* [ɔ] 'surgery'

Also: j*o* [ɔ] 'I', c*o* [ɔ] 'what', wj*o* [ɔ] (exclamation word used with horses), m*o* [ɔ] 'he has', st*o* [ɔ] 'hundred', ł*o* [ɔ] 'about', p*o* [ɔ] 'after', z*oo* [ɔ] 'zoo', dl*o* [ɔ] 'for', d*o* [ɔ] 'to', n*o* [ɔ] 'yes', pstr*o* [ɔ] 'nothing', ś*o* [ɔ] 'go away' ft*o* [ɔ] 'who'

A comparison of the data in (33) and (35) raises the fundamental question of whether there is any generalization to speak of. Six examples in (33) show final tense [o] and seventeen examples in (35) show final lax [ɔ]. There is a contradiction.

How do we know that the opposite of Final *ɔ*-Tensing is not true, that is, why can't we assume that final *o*'s in (35) are tense in the underlying representation and they become lax through the action of Final *o*-Laxing. Schematically:

(36) Final *o*-Laxing: o → ɔ / — #

The numbers alone, seventeen with final lax [ɔ] in (35) versus six with tense [o] in (33), give preference to Final *o*-Laxing over Final *ɔ*-Tensing. However, there are two reasons showing that Final *ɔ*-Tensing is a better analysis than Final *o*-Laxing: first, phonological plausibility and, second, the treatment of borrowings and foreign names.

Final *ɔ*-Tensing is a natural process supported, for example, by English. In *The Sound Pattern of English* (1968), Chomsky and Halle argue that words such as *motto* [mɑɾoʊ] and *veto* [viːɾoʊ] should derive their final diphthong by Final *ɔ*-Tensing, *ɔ* → *o*, and Diphthongization, *o* → *oʊ*. The argument is that *motto* could not have tense //o// and hence //oʊ// in the underlying representation because heavy syllables (here: syllables with a long vowel or a diphthong) attract stress:[30] SPE rules assign stress to the second syllable of *balloon*, *serene, suppose, retain*, and so forth. To conclude, *motto* derives from underlying //mɑtɔ//. The //ɔ// tenses

[29] This suffix is dying out in the form of [ɔ]. Our teenage consultants as well as those in their 30s and 40s confirm that the youngest generations of Goralians invariably realise the fem. nom.sg suffix as [a].

[30] The technical details are different, but they are not relevant. SPE used tenseness rather than length to define heavy syllables; see Chomsky and Halle (1968).

to [o] by Final Tensing and diphthongizes to [oʊ] by Diphthongization. Flapping completes the derivation, yielding the surface [mɑɾoʊ]. We conclude that Final *ɔ*-Tensing as a process is supported independently by stress facts of English.

The second argument for postulating Final *ɔ*-Tensing rather than Final *o*-Laxing comes from the behaviour of borrowings and foreign names.

(37)	Marcó Poló [marko pɔlo]	Galileó [galʲilɛo]
	Paló Altó [palo alto]	Toledó [tɔlɛdo]
	Tokió [tɔkʲjo]	Chicagó [ʃʲikago]
	discó poló [dʲisko pɔlo]	judó [ʤudo]

The final *o* is invariably tense in (37). The tensing must be due to a rule because the input has a lax [ɔ], as can be gleaned from the fact that lax *o* appears in the first syllable of *Poló* [pɔlo]. Borrowings and foreign names, unlike the morphemes in (33) and (35), constitute an open class, so it is not possible to assume that [o] comes from the underlying representation. We conclude that Final *ɔ*-Tensing is the correct descriptive generalization about Podhale Goralian.

The informal statement of Final *ɔ*-Tensing in (34) raises the question of how this generalization and the process it represents can be analyzed in terms of Optimality Theory (Prince and Smolensky 2004, McCarthy and Prince 1995). First, the rule needs to be restated as a constraint and, second, we need to identify faithfulness violations incurred by the tensing operation *ɔ* → *o*.

(38) Final *ɔ*-Tensing, **ɔ#*

No final lax [ɔ].

In terms of distinctive features, the tensing *ɔ* → *o* changes [-ATR] to [+ATR], which violates IDENT[±ATR].

(39) IDENT[±ATR]: The value of [±ATR] on the input vowel must be preserved on a correspondent of that vowel in the output.

The response to **ɔ#* can lead to a number of changes, apart from the desired *ɔ* → *o*. For example, final //ɔ// may change into [a] or [ɛ]. These undesired options are closed by faithfulness constraints, the relevant ones being the following.

(40) a. IDENT[±low]: The value of [±low] on the input vowel must be preserved on a correspondent of that vowel in the output.

b. IDENT[±back]: The value of [±back] on the input vowel must be preserved on a correspondent of that vowel in the output.

This discussion is summarized in (41), where we look at the evaluation of *Poló* of *Marcó Poló*.

(41) //pɔlɔ// → [pɔlo]

	*ɔ#	IDENT[±low]	IDENT[±back]	IDENT[±ATR]
(a) pɔlɔ	*!			
☞ (b) pɔlo				*
(c) pɔla		*!		
(d) pɔlɛ			*!	

Candidate (41a) loses because it has lax [ɔ] at the end of the word, a violation of *ɔ#. Candidates (41c) and (41d) violate the undominated faithfulness constraints and are hence *hors de combat*. Candidate (41b) violates IDENT[±ATR], but, crucially, this constraint is ranked below the other constraints in (41).

A successful analysis of Final *ɔ*-Tensing does not solve the problem of the data in (35), where final *o* is lax, as in *cȳt+o* [ɔ] 'he reads'. Most of these data are related synchronically or diachronically to forms with [a], as shown by the following comparison. Additionally, in *cȳt+o* [ɔ] 'he reads' – *cȳt+aj+om* [ɔ] 'they read', there is an alternation between [ɔ] and [a].

(42)	*Standard Polish*	*Kurpian*	*Podhale Goralian*	*gloss*
	czyt+a [a]	cyt+å [ɑ]	cȳt+o [ɔ]	'he reads'
	grub+a [a]	grub+å [ɑ]	hrub+o [ɔ]	'fat' (fem. nom.sg.)
	ja [a]	jå [ɑ]	jo [ɔ]	'I'
	dla [a]	dlå [ɑ]	dlo [ɔ]	'for'

We have identified the reason for the irregularity in the application of Final *ɔ*-Tensing: the [ɔ] from [a] does not undergo the rule. The question is how this descriptive/historical generalization can be incorporated into the phonology of Podhale Goralian. We have no answer to this question at this point.

6. Conclusion

The fieldwork carried out in Dzianisz has led to the discovery of new descriptive generalizations. First, the original *o* [ɔ] and the *o* from historical *a* that used to be different are now homophonous. Consequently, the word *Polok* 'Pole' from the earlier *Polåk* [ɑ] has two identical open *o* vowels: [pɔlɔk]. Second, *y* [ɨ], as in *ryba* 'fish' has shifted to tense *e*: [reba]. Third, Podhale Archaism is still distinct from the rest of Goralian phonology, but the distinction manifests itself in a different way: the 'archaic' *i* is now a central strongly retracted vowel [ɨ], so *cȳtać* 'read', *sȳn* 'son' and *zȳtó* 'corn' are pronounced [tsɨtatɛ], [sɨn] and [zɨto], respectively.

The fourth discovery is that *o* //ɔ// is tensed to [o] word-finally as in *wartkó* [vartko] 'quickly'.

In general, the vowel inventory of Podhale Goralian is different from that of Standard Polish in a way that is typical for Polish regional dialects: the system of mid vowels includes //e// and //o// in addition to //ɛ// and //ɔ//. The neutralization of the two *o*'s: the original *o* [ɔ] and the [ɔ] from [a] has disturbed the phonological system because the two *o*'s behave differently *vis-á-vis* phonological generalizations. Investigating this behaviour is an agenda for future research.

References

Avanesov, Ruben Ivanovič. 1968. *Russkoye literaturnoye proiznoshenie.* Moscow: Prosveshchenie.

Biedrzycki, Leszek. 1974. *Abriss der polnischen Phonetik*. Warszawa: Wiedza Powszechna.

Bilodid, Ivan Konstântinovič. 1969. *Suchasna ukraïns'ka literaturna mova. Fonetyka*. Kiev: Naukova Dumka.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English.* New York: Harper and Row.

Clements, George Nick and Samuel Jay Keyser. 1983. *CV Phonology. A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.

Decaux, Étienne. 1973. Tzw. archaizm podhalański a sprawa mazurzenia w Małopolsce. *Revue des études slaves* 49, 119-130.

Dejna, Karol. 1973. *Dialekty polskie*. Wrocław: Zakład Narodowy im. Ossolińskich.

Halle, Morris. 1992. Phonological features. In William Bright (ed.) *International Encyclopedia of Linguistics*. Vol. 3. New York & Oxford: Oxford University Press. 207-212.

Kąś, Józef. 2012. Podstawy ortografii podhalańskiej. In Krzysztof Kudłaciak (ed.), *Różnie tô na tym Bôzým świecie. Bukôwiańskie godki*. Bukowina Tatrzańska: Bukowiańskie Centrum Kultury „Dom Ludowy".

Kąś, Józef. 2015. *Ilustrowany leksykon gwary i kultury podhalańskiej.* T. 1. Bukowina Tatrzańska: Bukowińskie Centrum Kultury „Dom Ludow''.

Kopernicki, Izydor. 1875. *Spostrzeżenia nad właściwościami językowemi w mowie Górali Bieskidowych, z dodatkiem słowniczka wyrazów góralskich. Rozprawy i sprawozdania z posiedzeń Wydziału Filologicznego Akademii Umiejętności* III. Kraków: Drukarnia Uniwersytetu Jagiellońskiego pod zarządem Ignacego Stelcla.

Kosiński, Władysław. 1884. Przyczynek do gwary zakopiańskiej. *Rozprawy i sprawozdania z posiedzeń Wydziału Filologicznego Akademii Umiejętności* X. Kraków: Drukarnia Uniwersytetu Jagiellońskiego pod zarządem Ignacego Stelcla.

Kryński, Adam Antoni. 1884. Gwara zakopiańska. Studyjum dyjalektologiczne. *Rozprawy i sprawozdania z posiedzeń Wydziału Filologicznego Akademii Umiejętności* X. Kraków: Drukarnia Uniwersytetu Jagiellońskiego pod zarządem Ignacego Stelcla.

Małecki, Mieczysław. 1928. *Archaizm podhalański. Monografje polskich cech gwarowych* 4. Kraków: Akademia Umiejętności.

McCarthy, John J. and Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill N. Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.), *Papers in Optimality Theory.* Amherst: GLSA. 249-384.

Nitsch, Kazimierz. 1915. *Dialekty języka polskiego*. Kraków: Akademia Umiejętności. and 1957. 2nd edition. Wrocław: Zakład Narodowy im. Ossolińskich.

Prince, Alan and Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell. [Revision of 1993 technical report, Rutgers University Center for Cognitive Sciences. Available on Rutgers Optimality Archive, ROA-537.]

Rozwadowski, Jan Michał. 1915. *Historyczna fonetyka, czyli głosownia języka polskiego*. Reprinted in Jan Michał Rozwdowski, *Wybór pism. Tom I. Pisma polonistyczne.* 1959. Warszawa: Państwowe Wydawnictwo Naukowe.

Rubach, Jerzy. 2003. Polish palatalization in Derivational Optimality Theory. *Lingua* 113, 197–237.

Rubach, Jerzy. 2011. The vocalic system of Kurpian. *Studies in Polish Linguistics* 6:81-98.

Rubach, Jerzy. 2017. Derivational meanders of high vowel palatalization. *Lingua* 199:1-26.

Sagey, Elizabeth. 1986. *The Representation of Features and Relations in Non-linear Phonology.* Doctoral dissertation, MIT.

Schuster-Šewc, Hinc. 1968. *Gramatika hornjoserbskeje rěče*. Bautzen: Domowina Verlag. Translated into English by Gary H. Toops as *Grammar of the Upper Sorbian Language*. 1996. München-Newcastle: Lincom Europa.

Stieber, Zdzisław. 1952. *Rozwój fonologiczny języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.

Stieber, Zdzisław. 1973. *A Historical Phonology of the Polish Language*. Heidelberg: Carl Winter.

Stopka, Nazimek, Andrzej. 1897. *Sabała: portret, życiorys, bajki, powiastki, piosnki, melodye*. Kraków: Drukarnia Narodowa.

Stopka, Nazimek, Andrzej. 1911. *Rycerze spiący w Tatrach: fantazye na tle podań ludowych.* Kraków: Drukarnia i Zakład Art.-Litogr. Polonia.

POLISH MIGRANT COMMUNITY IN IRELAND: THE USE OF IRISH ENGLISH SLIT-T

IZABELA GRABARCZYK
University of Łódź, Poland
izabela.molinska@uni.lodz.pl

Abstract
In any migratory context individuals are faced with several challenges as a result of having to live in a different geographical location, function in a different cultural setting and use a different language. The migrants' use of language plays a crucial role in mediation of their identity, especially in the domain of pronunciation (Kobialka 2016). When non-native users of language adapt their speech to resemble that of the host community, it may suggest their strong identification with the target community (Hammer and Dewaele 2015). This papers focuses on the pronunciation patterns among Polish adult migrants living in the west of Ireland. The aim of the study is to investigate the link between positive attitudes of the migrant community towards Ireland, Irish culture and community, their acculturation strategies and language identity, and the tendency to use one of the most characteristic features of Irish English – slit-t. The theoretical framework includes acculturation theory (Berry 2005), social identity theory (Tajfel and Turner 1987) and language identity (Block 2007). The qualitative and quantitative analysis of data indicates a certain correlation between the use of Irish English slit-t and the participants' strategies of acculturation, identity and attitudes to the host community.

Keywords: migration, Polish migrants, second language use, acculturation, Irish English

1. Introduction and rationale

This project investigates pronunciation patterns among adult Polish immigrants living in the Republic of Ireland. The aim of the study was to explore the use of one of the most prominent features of Irish English: slit-t.

Polish migrant community is currently the largest non-Irish group in Ireland with population of almost 123 000 individuals (Census 2016). Compared to Polish diasporas in other countries, such as the United States or the United Kingdom, this is a relatively new community, one that is currently in the process of establishing itself. The largest influx of Polish migrants to Ireland occurred during the years following Poland's accession to the European Union in 2004. Prior to that, Ireland did not belong to primary destination countries for Polish migrants. The members of Polish community in Ireland were faced with a challenge of adapting to the new situation – as individuals, as members of a minority group and, most importantly to this paper, as users of English. Interestingly, for many newcomers from Poland, moving to Ireland involved their first encounter with Irish English, which is

decidedly different from varieties which may be encountered in Polish schools, especially in terms of accent. Furthermore, Polish nationals arriving in Ireland after the 2004 EU enlargement were not exposed to any patterns established by previous Polish migrant communities with regard to social or linguistic behaviours, since there were no patterns to follow.

In the light of the above numbers and circumstances, the way Poles use language in their everyday interactions with members of the host community or other migrants, allows insight into the relationship between language and identity, attitudes and acculturation of individuals in the host country. The main objective of this study was to investigate the link between these aspects – acculturation, identity and attitudes of Polish individuals towards Ireland, its people and culture, and the tendency for Poles to use characteristic features of Irish English pronunciation, namely Irish English slit-t, as an indicator of their level of integration into the Irish society and a possible marker of their newly acquired social or national identity. It is assumed that the more positive their attitude and the more successful their cultural adaptation, the more inclined they are to use local variants of pronunciation, which in turn may be an expression of accommodation, and aspiration to belong, to the target group.

2. Poles in Ireland

The years following the accession of Poland to the EU resulted in subsequent waves of Polish migrants moving to the UK and Ireland. Unlike Great Britain, Ireland did not belong to primary destination countries for Polish migrants before the 2004 EU enlargement. According to Grabowska (2005), there is some evidence indicating earlier migration flows of Poles to Ireland, although rather insignificant in numbers. The author described four waves of migration – post WW II, in the early eighties, in the mid-eighties and post-1997 during the Celtic Tiger boom years – the latter includes influx of Polish nationals after joining the European Union in 2004 (Grabowska 2005).

While the first three waves did not lead to any changes in the demographic composition of the Irish population, the phenomenon of the last wave was remarkable. It brought unprecedented numbers of Poles to Ireland. Since May 2004 all nationals of the EU Accession States were granted full access to the labour market in Ireland (Grabowska 2005: 31-32). The availability of employment opportunities served as one of the primary incentives for those planning to leave Poland in search of work to choose Ireland. What makes Polish diaspora in Ireland all the more interesting as a study subject is their extensive cultural and social activity, evident in numerous initiatives undertaken by Polish migrants, as well as organisations and associations aimed at bringing Polish migrants together and developing cooperation with the Irish society. Some of

these include Polish Irish Association, Irish Polish Society, PolskaÉire Festival or Polish weekend schools.

3. Acculturation and identity

This study draws on theories of acculturation, identity as expressed through language and attitudes – understood as the approach of immigrant community towards the variety of English spoken in Ireland, the host community and life in Ireland.

3.1. Acculturation

The concept of acculturation was first mentioned in the early 1930s. It derives from the field of cross-cultural psychology aimed at demonstrating the influence of cultural factors on the development of behaviour of individual humans. Cross-cultural psychology supports the view that individuals change their behaviour to suit the new setting (Berry 1997).

John W. Berry defines acculturation as "a process of cultural and psychological change that results from the continuing contact between people of different cultural background" (2006: 27). Every situation in which individuals brought up in two different cultures come into contact for an extended period of time facilitates acculturation (Waniek-Klimczak 2011). This phenomenon, commonly associated with migration, is understood as the process where individuals become adapted to a new culture, and their way of thinking and feeling undergoes reorientation (Brown 1994). There are differences between individuals and groups which influence the process of becoming members of new societies. As Berry notes "some jump in with both feet, seeking rapid absorption, while others are more hesitant, seeking to retain a clear sense of their own cultural heritage and identity" (2006: 30).

Berry (1997) proposed four strategies of acculturation, based on two components: individuals' attitudes (the preference) and behaviours (the actual practices) as displayed in daily intercultural interactions. The patterns of adaptation chosen by members of non-dominant groups are assessed on the "basis of bi-polar answers to two major questions about cultural maintenance on the one hand, and contact and participation on the other" (Waniek-Klimczak 2011: 229). These strategies include assimilation (members of the non-dominant group display negative attitudes towards maintenance of their native cultural identity and positive attitudes towards the culture of the host community); separation (members of the non-dominant group display positive attitude towards their native cultural identity and do not want to participate in the dominant culture); marginalisation (members of the non-dominant group display little interest in either, native culture maintenance and interaction with the host community); integration (members of the non-dominant groups display interest in both, native

culture maintenance and interaction with the members of the host community) (Berry 2006). The above classification of strategies is based on the assumption that the members of non-dominant group have the freedom of choice in terms of acculturation, although this may not always be possible (Berry 2006).

According to Waniek-Klimczak (2011), the integration strategy is most beneficial as regards individuals' general well-being in the long term, while assimilation strategy leads to successful language acquisition, which is the result of extensive exposure to target language (TL) and interaction with TL speakers. In Schumann's (1986) model of second language acquisition, acculturation is understood as "social and psychological integration of the learner with the target language group" and considered "a major causal variable in SLA" (1986: 379). This model is based on the prediction that the degree of an individual's acculturation into the target language community influences the degree to which an individual acquires the language of the target community (Hammer and Dewaele 2015). As Schumann (1986) notes, although acculturation is not the direct cause of SLA, it constitutes a crucial link in what he calls "a chain of causality" which leads to natural SLA (1986: 385). The level of acculturation determines the amount of interaction with the TL community, thus it influences the amount of input received by the SL speaker, which in turn directly influences the process of SLA. In his model Schumann (1986) included social and affective variables, which are believed to influence the quality and quantity of contact between SL users and TL community. One of the social variables which is of particular interest in this paper is the type of integration strategy, i.e. assimilation, preservation and adaptation, which correspond to Berry's strategies of assimilation, separation and integration respectively (Waniek-Klimczak 2011). In the assimilation strategy the contact between non-dominant and dominant language groups is enhanced, as the former group chooses not to maintain homeland values and instead "adopts those of the target language group" (Schumann 1986: 381). Preservation strategy refers to a reverse situation, i.e. non-dominant group chooses to maintain their values and reject the values of the target language group, which leads to an increased distance between the two and thus limited interaction and chances for acquisition of the language. Finally, the adaptation strategy refers to a situation in-between, the non-dominant group both maintains their cultural values within the group and, at the same time, adopts the values and lifestyle of the TL group, which has a varied influence on the language acquisition, depending on the degree of interaction (Schumann 1986).

3.2. Identity

The theory of social identity and language identity were demonstrated as influential factors in the acquisition and use of second language in the migrant context. According to David Block, "it is in the adult migrants' experience that identity and one's sense of self are most put on the line […] individuals are forced

to reconstruct and redefine themselves" (2007: 75). The reconstruction and redefinition process concerns both, the individual's sense of self, as well as the position they are assigned in the new environment (Block 2007). The naturalistic context does not necessarily guarantee sustained interaction with users of the second language (Block 2007). The roles individuals undertake as members of different social groups are of huge importance and so is their identity as expressed through language. This is particularly evident in the domain of pronunciation, which Lybeck (2002) described as the most prominent linguistic marker of individuals' cultural identification. Any changes in the pronunciation patterns to mirror native speakers is perceived as a distinct marker of cultural identification with the target language community and it may be viewed as manifestation of a new cultural identity being developed (Hammer and Dewaele 2015: 184).

In an article on language maintenance, Giles and Johnson (1987) drew on Social Identity Theory in the context of ethnic minorities:

> [P]eople who define an encounter with a member of another ethnic group as an intergroup one and value [the intergroup's] language as a core aspect of its identity will wish to assume a positive identity by means of adopting various strategies of 'psycholinguistic ethnic distinctiveness' such as switching to ingroup language, accentuating ethnic dialect and slang […] (Giles and Johnson 1987: 71)

Individuals use language in order to identify themselves with a given group to suggest that they belong to that group. This can be done in different ways, e.g. by adopting aspects of the language of that group, a specific dialect or an accent of a language, or by using slang or a given jargon, by employing language accentuation strategies, e.g. speech markers (Giles and Johnson 1987: 71). However, it may also happen that minorities which are characterized by negative social identity as compared to the dominant group, do not pursue those strategies (Giles and Johnson 1987). This may take place particularly when individuals "(a) strongly identify with their social group, and (b) make insecure social comparison between the positions of their group socially and that of the outgroup" (Giles and Johnson 1987: 71). Language is one of the attributes that is subject to huge impact of ingroup/outgroup dynamics.

Although Giles and Johnson (1987) focus primarily on the cognitive processes and behaviours in interethnic encounters, this approach can be easily applied in a migratory context. In the current study, the participants are members of a larger Polish community in Ireland, while the Irish host community constitutes the outgroup. Although this categorization is valid in general terms, it would be simplistic to consider it the only possible grouping. In order to be members of a given group, the participants must internalize the membership to that group as means of self-reference and it is not enough for others to classify them as such. Furthermore, individuals can belong to a variety of social groups at the same time and may more or less consciously adopt certain prototypical values, behaviours and norms characteristic of that group. Since language is one of the differences

between the groups in question, it may be affected and adopted in a variety of ways in order to achieve positive social identity and mark ingroup affiliation. Although there is no denying that becoming a member of a group influences the way individuals perceive themselves and may evoke a process of re-definition of self. The cognitive processes as well as behaviours vary between individuals, even if they are members of supposedly the same ingroup.

David Block also takes up the issue of second language identity, defined in terms of "degrees of audibility in the second language" (2007: 41). The concept of audibility may be described as "the right accent as well as the right social and cultural capital to be an accepted member of a community of practice" (Block 2007: 41). The process of developing an identity in L2 involves linguistic features, but also a range of semiotic behaviours, such as dress, expressions and movement (Block 2007). The expression of identity in second language is as complex and multivalent as it is in the case of one's native language. In a situation where individuals leave behind their home country and their native language, and they have to function in a new country, with a new culture and a new language, they might feel the need to re-establish their identity (Cook 2002). Whether consciously or without putting much thought into it, individuals may adopt different linguistic features to create their self-representation.

4. Irish English slit-t

Irish English slit-t is a fricated realisation of alveolar plosive /t/, known also as /t/ lenition. It is considered the clearest and the most extensive feature in the Republic of Ireland (Hickey 2008; Hickey 1986: 17). Although it can be found in other varieties of English (e.g. Liverpool English, Newfoundland English, Australian English), its origins were linked to Irish nationals who migrated to different parts of the world, taking their accent with them (Jones and Llamas 2008). The feature is present in supraregional Irish English, meaning it is not tied to any particular region of Ireland and may be encountered all over the country (Hickey 2012). However, while in supraregional Irish English t-lenition is always realised as apico-alveolar fricative, in local Dublin English other realisations may also found. The slit-t realisation occurs when preceded by a vowel and followed by another vowel or a pause (Hickey 1995). This allophone may be present in intervocalic (e.g. *sitting*) and word-final (e.g. *but*) position; it does not appear when preceded or followed immediately by another consonant sound (e.g. *backed*, *bootleg*) (Pandeli et al. 1997). The same process extends to voiced alveolar plosive /d/, although it is not as frequent and phonetically salient (Kallen 1997; Hickey 2007). In terms of the physical characteristics of the variable, Hickey (1996) emphasises the active articulatory role of the tip of the tongue, which distinguishes it from /s/ and /z/, the controlled manner of the movement and the lack of closure - the tongue

is held above the alveolar ridge, resulting in producing an apico-alveolar sound instead of a stop (Pandeli et al. 1997).

This variable has been chosen for several reasons. According to Hickey, Irish English slit-t as an allophone of /t/ is "one of the auditively most salient features of southern Irish English" (1995: 119). Furthermore, although there is some research on the slit-t as regards the native users of Irish English, the feature does not seem to have been studied in second language context, as far as the author is aware. Finally, the fricative realisation of alveolar plosives is not encountered in the Polish language, which excludes the possibility of language interference.

5. Methods

This project was an exploratory study of the relationship between the participants' acculturation patterns, identity, their attitudes towards the target language and the host community and the use of the Irish English slit-t. The sociolinguistic character of the study required the employment of a combination of quantitative and qualitative approaches in order to draw a clear picture of pronunciation patterns on one hand, but also to gain access to possible factors influencing these patterns on the other. The qualitative analysis of the semi-guided recorded interviews focused on the participants' background and the independent variables, i.e. language identity, their attitudes towards the language and the host community and the acculturation strategies. The quantitative analysis involved quantifying the number of occurrences of the dependent variable, i.e. Irish English slit-t.

6. Participants

The participants of the study are presented in the table below. There were ten adult Polish individuals living in the west of Ireland, eight from County Clare and two from the neighbouring County Galway. The sample was limited to the western seaboard of the country in order to exclude possible regional variation. The participants were recruited via the friend-of-a-friend sampling technique (Milroy 1992), that is participants of the study assisted the researcher in recruiting other prospective participants, e.g. their friends, neighbours, acquaintances. Eight of the participants moved to the Republic of Ireland in the post-2004 accession while two of them moved to Ireland shortly before that date (the exact number of years is listed in the table below under Length of Residence). All of the participants were in full or part-time employment in a variety of professions, including IT, retail, engineering, manufacturing, customer services and accounting. The sample includes five male and five female participants in order to minimize possible gender differences. The age of the participants ranges from 31 to 40 years of age. Due to the limited length of the paper, the participants' profiles can be found in the *Appendix* section.

Table 1. Participants

Participants	Age	Sex	Resides in	LoR (years)
Speaker A	32	Female	Ennis	8
Speaker B	32	Female	Ennistymon	10
Speaker C	39	Male	Ennis	10
Speaker D	32	Female	Ennis	9
Speaker E	31	Male	Ennis	10
Speaker F	38	Female	Ennis	10
Speaker G	32	Male	Galway	12
Speaker H	40	Male	Ennis	12
Speaker I	37	Male	Ennis	11
Speaker J	40	Male	Galway	13

7. Data collection

Data collection was implemented in person. The interviews were conducted in English, either in the researcher's or the participant's house and took between 15 minutes and 2 hours. The participants were made aware beforehand that the interview requires a level of English which is sufficient to hold a conversation in English for a longer period of time. The first part of the project included a semi-guided interview during which a number of issues concerning the participants' immigration experience in Ireland were addressed. The aim of the interview was to elicit a variety of linguistic data both in formal and informal context via a number of strategies based on Labov's (1984) sociolinguistic interview. The researcher focused on spontaneous speech (as far as possible), including a set of previously prepared guiding questions aimed at evoking emotional response and creating a friendly atmosphere, but also to elicit necessary information concerning the participants' life in Ireland, their migratory experience, attitudes towards the host community and the target language, their social and national identity, as well as possible acculturation strategies. In the second part of the meeting, the participants were given samples of English to read - a number of sentences and a text passage containing the dependent variable. The data was recorded with the use of a H1 Handy Recorder and the interviews were recorded in .wav format, using 44.1 kHz sampling rate.

8. **Data analysis**

The sample recordings were processed for auditory analysis in Audacity, relevant data was later analysed in Praat, where spectrograms were generated in order to support the auditory judgement. The qualitative analysis included a review of the answers to the questions and comments given by the participants in the semi-guided interview. The questions/guiding topics may be roughly organized into the following categories: migratory experience; awareness of Irish English as a variety and of variation within the country; attitudes towards preserving Irish/Polish customs and language; attitudes towards Irish/Polish community (and awareness of the prominence of the latter); active involvement in the life of Irish/Polish community; participants' future plans; participants' favourite places, likes and dislikes as regards their home country and the host community, ways of spending their free time.

The quantitative part of the analysis involved the analysis of the sentences containing the target variable, as well as data gathered during the semi-guided interview. The nine carrier sentences included four instances of alveolar plosive /t/ in word-medial and five instances of /t/ in word-final position, which creates context for realising the sound as slit-t. The data was subject to auditory analysis in Audacity and a selective visual analysis of spectrograms generated in Praat software in order to assess the closure or a lack of closure (see Figure 1 below) in potential instances of the target sound. All of the instances of the alveolar plosive /t/ realised as fricated variants were marked as 1, while other realisations of the sound were marked as 0. As regards the data gathered during the semi-guided interviews, the following procedure was adopted: ten first utterances containing /t/ sound in relevant environment (i.e. intervocalic word-medial and word-final position, where the alveolar plosive could possibly be realised as Irish English slit-t) were extracted and subjected to auditory as well as visual analysis. Similarly to carrier sentences, all utterances where fricated realisation of /t/ occurred were marked as 1, while other realisations were marked as 0.

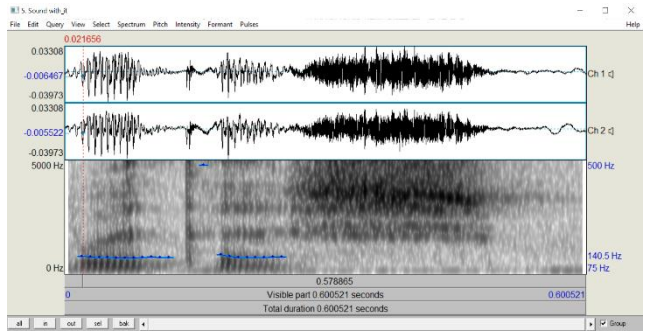


Figure 1. Example of no closure in ‘with it’.

9. Results

The results section gives an overview of quantitative analysis, i.e. whether and to what extent the participants adapt the feature in question. The profile of each participant based on the answers to interview questions can be found in the *Appendix* section. The descriptions of individual participants allow for a better understanding of the speakers' background as regard their migratory and language experience, and hence the possible influences on their speech patterns.

This section contains both the results of the quantitative analysis of the carrier sentences (controlled speech) and the utterances extracted from the semi-guided interviews (spontaneous speech). As described in the previous section, all the instances of /t/ realised as slit-t were marked as 1, other realisations were marked as 0. The analysis was based on auditory impression, supported by spectrograms generated in Praat. The quantitative approach is justified by the exploratory character of the study and the author's aim to investigate the tendency to incorporate slit-t by Polish users of English. The results for carrier sentences are presented in *Table 2* and for utterances in *Table 3*.

Table 2. Instances of Irish English slit-t in carrier sentences

Participants	**Instances of Irish English slit-t**									
	butter	*letter*	*meter*	*water*	*but*	*what*	*hat*	*feet*	*right*	**Total**
Speaker A	1	1	1	1	1	1	1	0	1	**8**
Speaker B	1	1	1	1	1	1	0	0	0	**6**
Speaker C	1	1	1	1	1	0	1	0	1	**7**
Speaker D	1	1	1	1	0	1	0	0	1	**6**
Speaker E	0	0	0	0	1	1	1	0	0	**3**
Speaker F	1	1	1	0	1	1	0	0	0	**5**
Speaker G	0	0	0	0	0	0	0	0	0	**0**
Speaker H	1	0	1	1	0	0	0	0	0	**3**
Speaker I	0	0	0	0	0	0	0	0	0	**0**
Speaker J	0	0	0	0	0	0	0	0	0	**0**
Total	**6**	**5**	**6**	**5**	**5**	**5**	**3**	**0**	**3**	

Table 3. Instances of Irish English slit-t in the interview

Participants	Instances of Irish English slit-t										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**	**Total**
Speaker A	1	1	1	1	1	0	1	1	0	1	**8**
Speaker B	0	0	0	0	0	0	1	1	1	1	**4**
Speaker C	1	1	0	1	1	1	1	1	0	0	**7**
Speaker D	1	1	1	0	0	1	0	1	1	0	**6**
Speaker E	1	0	1	0	1	1	0	0	1	1	**6**
Speaker F	1	1	0	0	1	1	1	1	1	1	**8**
Speaker G	0	1	0	1	0	0	0	0	1	0	**3**
Speaker H	0	0	0	0	0	0	0	0	0	0	**0**
Speaker I	0	0	0	0	0	0	0	0	0	0	**0**
Speaker J	0	0	0	0	0	0	0	0	0	0	**0**

In carrier sentences, *Speaker A* demonstrated systematic t-lenition, using slit-t realisation in eight out of nine possible contexts, both in word-medial environment, e.g. *butter*, *letter*, *meter, water*, as well as in word-final position, e.g. *hat* or *but,* although not in the word *feet*. As regards the spontaneous speech, *Speaker A* consistently used the lenited variant of alveolar plosive /t/, pronouncing the sound in question in eight out of ten instances, especially in high-frequency one-syllable function words, such as *but*, *it*, *that*, but also in content words such as *write* and *cute*, but not in *mate*. *Speakers B* and *D* produced slit-t in six out of nine instances in carrier sentences, while in the utterances *Speaker B* pronounced the lenited variant only in four instances out of ten, e.g. in *chat* and three times in *but*, while *D* had two more occurrences, similarly to *Speaker A*. *Speaker C* was also quite consistent in producing the target sound (eight out of nine instances), pronouncing a stop variant only in *what* and *feet*. In the utterances *Speaker C* produced the variable in question in seven out of ten contexts. In the sample by *Speaker E*, there were only three instances of slit-*t* in the controlled environment and six instances in the spontaneous speech. *Speaker F* displayed a tendency towards quite a consistent use of the variable in the carrier sentences (five out of nine) and even more so in the case of utterances, where in eight out of ten cases the sound was realised as slit-t. *Speakers G*, *H*, *I* and *J* displayed similar pronunciation patterns. In the carrier sentences, only *Speaker H* pronounced slit-t in two instances, while in the utterances *G* pronounced the variable in three contexts. In the controlled speech sample there was a general tendency among the informants to use the target sound in word-medial rather than in the word final position. The remaining participants did not display any trace of the variable in their speech.

10. Analysis and discussion

According to the background literature, identity, attitudes towards the host country and the target language, as well as acculturation strategies, play a significant role in the use of L2 phonology. As has been discussed, patterns of acculturation chosen by migrant users of L2 aid in predicting whether individuals will succeed in L2 acquisition and to what extent (Waniek-Klimczak 2011). Therefore, the participants who lean towards the assimilation strategy of acculturation should also demonstrate a greater degree of the use of the new, assimilatory target pronunciation. Similarly, the participants who express positive attitudes towards the Irish community and the Irish English language, should also display a higher degree of use of L2 phonology than the participants who do not share the same views on the host country. What is more, "the acquisition of local speech features could be viewed [...] as indicative of a growing sense of local identity" (Drummond 2012: 112). Therefore, the more favourable perception of the host community and the value attributed to the language, the more likely are the participants to identify with the dominant group, which may be displayed in the adoption of language accentuation strategies to mark their association with that group (Giles and Johnson 1987: 71).

The results indicated that the participants who expressed a positive attitude towards the host group and the target language, the ones who actively took part in the life of the Irish community, also demonstrated a higher level of Irish English slit-t use. Regarding the strategies of acculturation, it becomes fairly apparent that a clear-cut categorisation of participants is not possible as individual informants demonstrated quite contradictory attitudes. All of the participants showed tendencies towards either the assimilation or integration strategy, none of them demonstrated signs of marginalisation or separation. *Speakers A, C* and *D*, who have very limited contact with the members of the Polish community and interact primarily with Irish nationals at work, at home and in their social life, display a pattern of behaviour that may be described as an assimilation strategy (although their attitudes towards the Polish community do not appear negative, rather indifferent). The remaining speakers often interact with both communities equally, they usually work with Irish nationals but share accommodation with other members of the Polish community, spend time with other Poles or have a family with whom they communicate mainly in Polish. In terms of attitudes, the majority of the participants believe that preserving the culture and language of both communities is important. There are some exceptions, such as *Speaker A*, who does not seem to feel the need to cultivate Polish customs, or *Speaker I* who emphasises that celebrating the customs of the host community is important, but Polish nationals should not "pretend" to enjoy them in the same way Irish nationals do, or *Speaker H* whose family favours the traditions of the host country. The majority of the participants state the importance of cultivating the traditions of both countries, although in practice a lot of them appear to choose to celebrate

the customs they enjoy or believe to be important for some reason, regardless of whether these involve Polish or Irish traditions. It becomes evident yet again that the issue of acculturation is quite complex. Assigning participants to one strategy seems rather impossible and there are a lot of discrepancies between what the participants say and what they actually do.

With respect to social and national identity, seven participants claimed not to feel the need to be recognised as Polish nationals, while only three of them admitted they are "proud to be Polish." *Speaker C* admitted to "hiding" his nationality in the past in order to avoid negative comments on the part of the host community and, although his approach has changed in the recent years, he still avoids revealing his origins unless necessary. *Speaker I* claimed to be quite weary of flaunting his nationality as he does not want to be treated differently because of his origins, implying experiences that are rather negative – which he claims to be often the case in Ireland. The participants avoided identifying or being identified with the Polish community, because they either perceived it negatively or believed that it is negatively perceived by other groups. The author's general impression during the interviews was that the participants either claimed to be proud of their Polish origin right away or stated that they were not bothered about being recognised as a person of Polish origin, because that is simply what they were. Some of the participants were initially reluctant to answer the question, claiming that they do not know or have not given it a thought. However, the attitudes of all the participants towards the country, the Irish community and Irish English variety are positive. They unanimously defined Irish people and what they describe as their relaxed, easy-going way of life as the best part of their experience in Ireland. Speakers mention the friendliness and hospitality of the host community, and some emphasise similarities between Polish and Irish nationals as the source of positive mutual attitude.

All of the participants claim to be fond of the variety of English spoken in Ireland, some prefer it to "English English," some note it is easy to learn and understand once used to the accent, some participants even claim to "love it." Thus positive attitudes reflected in the use of language might be expected. As can be inferred from the quantitative and qualitative results, the participants who demonstrated the highest degree of consistency in the use of the variable in question include those who display traits of assimilation and integration strategies of acculturation, e.g. *Speaker A* and *C*. This might be the result of extended interaction and complete natural immersion in the target language. What is more, both participants *A* and *C*, take pride in being praised for their "Irish accent," which may be a case of "passing for a native," noted by Drummond (2012). Similarly, *Speaker B* displays quite consistent use of the target feature in the carrier sentences, which may reflect naturalistic language learning processes in a close-knit rural community where she "imitated the locals", who welcomed her and treated her as one of their own. What is interesting, however, is that *Speakers B*, *E* and *F*, spent a large portion of their day speaking English, but they also interact in Polish with their housemates, family or partners. Similarly, *Speakers*

G, *H*, *I* and *J* also claim to speak English for a considerable part of their day, yet they have very little or no trace of slit-t. Interestingly, all of them claimed that good pronunciation is important only if its lack hinders communication. *Speakers I* and *J* both stated that they were aware of their "Eastern European" accents, however while *J* claimed not to be bothered about it, *Speaker I* did not seem to be equally at ease in that matter. The lack of Irish English slit-t in the participants' speech may indicate that they are either not aware of the variable as a salient feature in the variety of English spoken in Ireland or they may deliberately stick to their pronunciation as a sign of being content with their identity or accepting it as something that cannot be altered.

It is also interesting to notice that several participants who demonstrated the use of the variable had a tendency to exaggerate the feature or use it in contexts where it would not normally appear. In words such as *Ireland*, *different*, *accent* several speakers used a heavily fricated or affricated variant of /d/ and /t/. Auditory analysis indicated that some of the participants have the tendency to overextend fricated realisation of alveolar plosive /t/ to other environments where it would be normally blocked by the phonetic environment, i.e. a preceding sonorant or a following plosive. However, visual analysis of the spectrograms generated in Praat indicated that a closure was present thus suggesting a possible affricated variant of the sound. Contrary to what has been noted in the previous paragraph, this realisation may suggest that the participants are aware of the salience of the feature after all and appropriate it as a prestigious variant used by the target group, the variety to strive for, but the level of awareness does not allow them to imitate the exact use of the feature by the host community.

11. Conclusion

Certain tendencies in the relationship of identity, attitudes, acculturation and the use of Irish English slit-t can be noted. However, while some aspects of the migration experience in Ireland influence the use of the local phonological feature by some participants, for others they did not play any role.

Further research might explore several other issues. First of all, it has been already noted that the assessment of acculturation strategies is rather troublesome, as the participants express contradictory statements regarding their relationship with the traditions of the home and host countries. The attitudes that can be derived from the statements very often do not match the participants' actual behaviour and practices. Furthermore, it became rather evident during the analysis of the data that the nature of interaction between Polish migrants and the TL community played a significant role. Those participants who demonstrated the highest degree of slit-t use seemed to have much closer relationships with the members of the host community. They participated in intimate celebrations, e.g. weddings, funerals, they spent their free time with Irish people, they were invited

to their homes and included in typically Irish festivities. On the other hand, the participants who displayed very little or no trace of the target feature interacted with the Irish community on a daily basis, but the private and intimate activities in their lives included mostly other Polish nationals. What is more, while some of the recipients took pride in being praised for their Irish accent and in being accepted by the host community, others did not feel the need to negotiate their identity to function in the target language community at all.

It appears that the issue is more complex and requires further analysis on a much larger scale. The sample in this study was relatively small and did not allow for a broader analysis of the relationship between identity, attitudes, acculturation and the pronunciation patterns of the Polish community in Ireland. A future study should include a much larger sample from informants across the country.

References

Berry, John Widdup. 1997. Immigration, Acculturation and Adaptation. *Applied Psychology: an International Review* 46.1. 5-68.

Berry, John Widdup. 2006. Contexts of acculturation. In David Lackland Sam and John Widdup Berry (eds.), *The Cambridge Handbook of Acculturation Psychology,* 24-42.Cambridge: Cambridge University Press.

Block, David. 2007. *Second language identities*. London: Continuum International Publishing Group.

Brown, Douglas. 1994. *Principles of language learning and teaching*. New Jersey: Prentice Hall Regents.

Central Statistics Office. 2016. This is Ireland: highlights from census 2011, part I. [Online] Available from: https://www.cso.ie/en/csolatestnews/presspages/2017/census2016summaryresults-part1/ [Accessed on: 27 May 2016]

Cook, Vivian. 2002. *Portraits of the L2 User*. Clevedon: Multilingual Matters.

Diskin, Chloé. 2016. Standard language ideologies in multicultural Ireland: A Case Study of Polish and Chinese Migrants in Dublin. In Vera Regan, Chloé Diskin and Jennifer Martyn (eds.), *Language, Identity and Migration,* 287-325.Oxford: Peter Lang.

Drummond, Rob. 2012. Aspects of identity in second language: ING variation in the speech of Polish migrants living in Manchester, UK. *Language Variation and Change* 24. 107-133.

Giles, Howard and Johnson, Patricia. 1987. Ethnolinguistic identity theory; a social psychological approach to language maintenance. *International Journal of the Sociology of Language* 68. 69-99.

Grabowska, Izabela. 2005. Changes in the international mobility of labour; job migration of Polish nationals to Ireland. *Irish Journal of Sociology* 14(1). 27-44.

Hammer, Kate and Jean Marc Dewaele. 2015. Acculturation as the key to the *ultimate* attainment? The case of Poles in the UK. In Fanny Forsberg Lundell and Inge Bartning (eds.), *Cultural Migrants and Optimal Language Acquisition*. Bristol: Multilingual Matters: 178-202.

Hickey, Raymond. 2012. Rural and urban Ireland: A question of language? In Irene Gilsenan Nordin and Carmen Zamorano Llena (eds.) *Urban and Rural Landscapes in Modern Ireland: Language, Literature and Culture*. Oxford: Peter Lang: 17-38.

Hickey, Raymond. 2008. Feature loss in the 19th century Irish English. In Terttu Nevalainen, Irma Taavitsainen, Päivi Pahta and Minna Korhonen (eds.), *The Dynamics of Linguistic Variation:*

Corpus Evidence on English Past and Present, 229-243.Amsterdam: John Benjamins Publishing.
Hickey, Raymond. 2007. *Irish English: history and present-day forms*. Cambridge: Cambridge University Press.
Hickey, Raymond. 1996. Lenition in Irish English. *Papers from the International Conference on Language in Ireland. Belfast Working Papers in Language and Linguistics* 13. 173-193.
Hickey, Raymond. 1995. An assessment of language contact in the development of Irish English. In Jacek Fisiak (ed.), *Language Change under Contact Conditions*. Berlin: Mouton deGruyter. 109-130.
Hickey, Raymond. 1986. Possible phonological parallels between Irish and Irish English. *English World-Wide* 7(1). 1-21.
Jones, Mark Jonathan and Carmen Llamas. 2008. Fricated realisations of /t/ in Dublin and Middlesbrough English: An acoustic study of plosive frication rates and surface fricative contrasts. *English Language and Linguistics* 12. 419-443.
Kallen, Jeffrey. 1997. Irish English: context and contact. In Jeffrey Kallen (ed.) *Focus on Ireland,* 1-33. Amsterdam: John Benjamins.
Labov, William. 1984. Research methods of the project on linguistic change and variation. In John Baugh and Joel Sherzer, (eds.), *Language in use: Readings in sociolinguistics,* 28-53. Englewood Cliffs: Prentice Hall.
Lybeck, Karen. 2002. Cultural identification and second language pronunciation of Americans in Norway. *The Modern Language Journ*al 86(2). 174-191.
Milroy, James. 1992. *Linguistic variation and change*. Oxford: Blackwell.
Pandeli, Helen, Eska, Joseph Francis, Ball, Martin John and Joan Rahilly. 1997. Problems of phonetic transcription: the case of the Hiberno-English slit-t. *Journal of the International Phonetic Association* 27. 65-75.
Schumann, John. 1986. Research on the acculturation model for second language acquisition. *Journal of Multilingual and Multicultural Development* 7(5). 379-392.
Waniek-Klimczak, Ewa. 2011. Acculturation strategy and language experience in expert SLA speakers: An exploratory study. *Studies in Second Language Learning and Teaching* Vol 1(2). 227-246.

Appendix

Participants' profiles

This section briefly describes the participants based on the information collected during the interviews as well as from other sources (private conversations, communication via social media).

Speaker A is a 32-year-old female from Ennis in County Clare working in retail. Prior to arriving in Ireland, she had been learning English for ten years, although she claims her passive knowledge of the language was very good and her speaking abilities were rather poor. In Ireland she learned English via interaction with the locals. She rates her language skills quite highly and expresses fondness of the variety of English spoken in Ireland. She claims to speak only English on everyday basis, with her Irish partner and co-workers. She is aware of different English accents but she does not find Irish English difficult. She is confident that she does

not have a Polish accent, but rather "some bit of a Clare accent." The participant notes that Poland "doesn't feel like 'home home' anymore, nice to see the family but it is more like a holiday than home." She notes that she understands why Polish people might want to "get this Polish tradition thing going on", but she does not express membership to any Polish groups or associations.

Speaker B is a 32-year-old female living in a small coastal town of Ennistymon in the west of County Clare. She works as a shop assistant for a small, family-run business. She is a single mother of three children. Her knowledge of English was fairly limited prior to her move to Ireland and she also learned English via interaction with her Irish and multi-national co- workers, as well as "chatty" customers, members of a close-knit community in a rural area. She pinpoints Irish English accent as the most difficult part of the language experience in Ireland, although she finds it easier to understand than American and British varieties. She notes that she has been told several times that she "must be from Clare" due to her accent. She uses the language mostly at work, while at home she communicates in Polish.

Speaker C is a male in his late thirties, he has tertiary level education and is working as an engineer in the IT sector. He received formal instruction in English and also spent several months in America and Spain prior to arrival in Ireland. English plays a primary role in his everyday life and he sometimes does not use Polish for longer periods of time. He refers to Irish English accent as "this is the killer" and describes it as one of the most difficult aspect of the language experience in Ireland. He believes that English is not necessary to succeed in Ireland as it depends on one's profession and as long as pronunciation does not hinder communication, it is not that significant. He also expressed his regret about initially "hiding" his Polish nationality. He believes that preserving both Polish and Irish customs is equally important. He owns a house in Ireland and is planning to stay, considering it a good place for living.

Speaker D is a female who works in a manufacturing company in Ennis. She knew little English prior to her arrival and learned it mostly in Ireland, also via interaction with the host community and her Irish partner. She expressed fondness of the Irish English variety but also stressed the difficulties connected with pronunciation, especially in rural areas. She identifies the ability to speak English as one of the crucial aspects of being able to succeed, although she notes that there is "no harm" in being able to speak Polish. The emphasized the importance of contributing to and socializing with the host community. She obtained Irish citizenship.

Speaker E is a 31-year-old male who moved to Ireland in 2006. He is employed in the services sector in the accounting department. He did not receive any formal instruction in English prior to his arrival and learned the language by watching movies and cartoons. He rates his English language abilities as very good, in fact he claims to be often taken for an Irish citizens and his interlocutors react with surprise once they learn his nationality. The speaker implied it might be due to his appearance, noting that people tend to ask him: "are you Irish-

looking speaking Polish or Polish-looking speaking English well." He is "proud to be Polish" and believes that preservation of the customs is very important which he practices with his Polish housemates. Irish customs are equally important to him. He works in an all-Irish environment and has a lot of Irish friends, hence English is the language of communication for a major part of his day. He considers speaking Polish "a massive advantage" in Ireland and the knowledge of this language sometimes helps him at work when his Irish co-workers ask for his assistance with Polish customers. He notes that the Irish accent, especially in rural areas, is the most challenging aspect of the language experience, although he believes that pronunciation is important as it might hinder communication. He is considering moving to a different country to "try something new."

Speaker F is a female in her late thirties. She has tertiary level education which she received in Poland and is currently working in a manufacturing company as an engineer. She received formal instruction in English in primary and secondary school, and at the university. She also speaks some Spanish, Swedish and Czech, the latter is also the language of her husband. She uses English every day, at work and with her friends. She considers Irish English accent a potential difficulty, although she is fond of this variety, claiming that she prefers "English [spoken] here [to] that in the UK." She believes knowledge of English is very important in order to achieve anything in Ireland and so is the pronunciation. She notes that it is important to preserve the Polish language, although as she notes it may be both an advantage and disadvantage - she does not elaborate on this claim, only comments "we are foreign here right." The speaker is aware of the Polish community in her town, but she is "not a member of that community." The participant is fond of her current place of residence, however, she considers a possible move in the future, but not to Poland.

Speaker G is a 32-year-old male living in Galway city. He arrived in Ireland two years prior to the accession of Poland to the EU. He is currently working in a technical support sector. Prior to coming to Ireland he learned English for five years, although he did not find that knowledge very useful as it was mainly "grammar-oriented" and he did not know how to "use" the language. Once in Ireland, he did a six-month language course in Galway and Galway Community College. He also pursued third level education in Ireland and holds a bachelor's degree from an Irish university. He uses a lot of English on a daily basis, especially with his co-workers and customers. Although the participants claims to know a lot of Irish nationals, he notes that "for some reason" he does not have any close Irish friends and majority of his closest friends are Polish. The participant believes that preserving the language and customs of both countries is crucial, especially for families with children, as it is the parents' responsibility to teach children Polish language and about the history of the country - one should "know their roots." He also strongly believed that the Polish community should learn about Irish history and traditions and "assimilate with the Irish people" by cultivating those customs. He admits to be proud of being Polish and considers himself a

member of the general Polish community living in Ireland, he has never been "ashamed" or tried to hide his origins. He is currently house-hunting and plans on settling in Ireland long-term.

Speaker H is a 40-year-old male, who had been living in Ireland for twelve years at the time of the interview. He works in a manufacturing company. Prior to settling in Ennis, he spent a year and a half in Cork, a city located in the south-west Ireland. He took several English courses in Ireland and he currently holds a FETAC TESOL certificate in English. Contrary to previous speakers, he does not consider Irish accent difficult, quite the opposite, he claims it is easy to understand despite the differences and he likes the way it sounds. He uses English at work with his co-workers and notes that knowledge of English is necessary and he would not be where he is without having sufficient language skills. He is confident in his English skills, although he notes that native English speakers can immediately tell that he is a foreigner. He considers himself a member of the Polish community in his town and occasionally participates in events organised by them. Preservation of Polish customs and language is important to him, as he has two daughters and believes it necessary for them to know their parents' country of origin, its language and history. He believes that speaking Polish is an advantage and facilitates a lot of interaction in a country with a large Polish community. He is proud of his origins and considers it important to be recognised as a person of Polish nationality. His children attend an Irish-medium school and are able to use both English and Irish, as well as Polish. He notes that preserving Irish and Polish customs is equally important. He visits Poland a couple of times a year, as he claims, he "doesn't like Poland as a country for living." He is a holder of an Irish passport and plans on staying in Ireland.

Speaker I is a male in his late thirties. He is currently employed in manufacturing. He had some formal instruction in English at a beginner level during his university years back in Poland, hence he notes his ability to speak English was limited upon arrival. He did not attend any English school while in Ireland, but rather "immersed" himself in language through books and television. He does not feel the need to read in Polish, or any language other than English, due to the wide spectrum of data available in English. The participant uses English on a daily basis at work, he also speaks English with his daughter six days a week, while one day is reserved for Polish as a way of "keeping [the daughter] bilingual." He also claims he pays a lot of attention to pronunciation, as he does not want to be a bad model for his daughter. He claims, however, that at this stage his daughter's English skills begin to surpass his knowledge of the language. The participant states he is aware of his strong eastern European accent, which he finds quite challenging. The speaker believes that his accent and nationality are the reasons for being approached in a different manner, not necessarily positive, by the host community. He states he is a member of the Polish community in Ireland "by birth," although he does not associate himself with them and does not "force" himself to interact with other Polish nationals. Quoting the speaker: "just because we are all Poles, doesn't mean we have to all meet." The speaker believes that it

is important for Polish nationals to learn about Irish history and culture and expresses his disapproval of certain members of the Polish community who overly demonstrate their patriotism without having basic knowledge of Polish history. The speaker stresses the importance of celebrating Irish customs, although without pretending "that we enjoy and appreciate them the same way as locals - we should respect them." He visits Poland about once a year and notes that "it doesn't feel like home at all."

Speaker J is a 40-year-old female living in Galway. She arrived in Ireland in 2003. Prior to her arrival, she had English classes in secondary school. Once in Ireland, she enrolled into a three-month-long business English course. The speaker works in an English-only environment, hence it is the only language of communication for a majority of her day. At home she speaks Polish with her partner. The participant states that although accent may pose some difficulty, she does not find it challenging after so many years in Ireland. She emphasises the importance of speaking English if one wants to accomplish anything in Ireland and claims she "loves" Irish English and "hates" the way English sounds in England. She believes she has a strong Eastern European accent, but as long as people understand her, it does not trouble her. She does not belong to any Polish or Irish associations. She believed it is natural for Polish people to speak Polish and preserve the customs. She notes that her nationality influenced her experience in Ireland in a positive way, as "Irish people find Polish people very similar to them in nature," they understand Poles' need to emigrate in order to improve the standard of their lives. She believes that the warm welcome she received from the host community would not be the case if she were of a different nationality. At the same time she notes it is not necessarily important to her to be recognised as a Polish national. In terms of traditions, the participant claims she "goes with the flow" and celebrates the customs of both countries. She is a holder of an Irish passport and is planning on staying in Ireland for now, although this may change in the future.

POSTER, POSTER ON THE *WALL*, DO YOU REALLY MEAN IT ALL?

DECODING VISUAL METAPHOR '*GLOBAL WARMING'* IN PUBLIC AWARENESS CAMPAIGNS

MARINA PLATONOVA
Riga Technical University, Latvia
marina.platonova@rtu.lv

Abstract
The tendency to create messages using the elements belonging to different semiotic systems shifts our perception of a communicative act, contributing to the establishment of multimodal and intersemiotic communication practice.
A visual metaphor is seen as one of the instances of a multimodal and intersemiotic message, which generates a text that is revealed gradually, uncovering numerous layers of meaning encoded within a metaphor and within visual, linguistic, and spatial settings it is placed in.
The paper sets out to explore the notion of a visual metaphor and focuses on the application of the visual metaphor '*global warming'* on posters created for the needs of public awareness campaigns, investigating simultaneous manifestation of iconic and metaphorical mappings in the given visual metaphor.

Keywords: environmental communication, iconicity, intersemiotic communication, intermediality, poster analysis, public awareness campaign, semiotic systems, visual metaphor

1. Introduction

In the contemporary knowledge society that "promotes multiple forms of semiotic systems" (Iyer and Luke 2010: 21), the process of information transfer calls for extensive multiliteracy on the part of the message recipients especially in the dynamically developing research fields.

The emerging interdisciplinary field of environmental communication represents the combination of special codes typical of not only distinct scientific domains, but even of different symbolic systems, as it communicates meaning "through the system of common symbols, signs and behavior" (MWD), thus forming a unified '*symbolic medium'* (Cox and Pezzullo 2015: 20) (see Platonova 2015 for discussion). This tendency shifts our perception of a communicative act, contributing to the establishment of multimodal and intersemiotic communication

practice, where the former concerns multiplicity of communication modes, while the latter addresses multiplicity and interrelation of semiotic systems.

The openness of a semiotic system is determined by its ability to co-exist with other semiotic systems, as well as by the ability of a user to recognise and make possible use of its constituent elements inviting new meaning combinations and "creating multiple moments of interpretive possibility" (Albers and Murphy 2009: 4). Interpretation implies intersemiotic translation, especially for the multimodal messages, where verbal signs are read by means of non-verbal signs and vice versa (cf. Petrilli 2003: 18). The combination of elements from different semiotic systems requires a recipient to establish comprehensible intertextual and intermedial links to the previously acquired set of knowledge to grasp the meaning of a message, which can be displayed both verbally and visually. The communicated message is defined as "meaningful combination of terms in a semiotic system" (Purchase and Naumann: 2014), whereas; "a physical realization of a message" is seen as a text (*ibid*), hence, following Albers and Murphy (2009), all forms of semiotic systems in use create a text.

A visual metaphor is seen as one of the instances of a multimodal and intersemiotic message, which generates a text, a whole story being narrated to the recipients. It is revealed gradually uncovering numerous layers of meaning encoded within the very metaphor and within the whole visual, linguistic, and spatial settings it is placed in. The blending of different semiotic systems for metaphorical meaning transfer calls for a convenient analytical tool, which would explain the mechanisms of concept cognition (encoding, decoding), metaphorical mapping, and imagery exploited for message communication. Therefore, to ensure unambiguous communication of information provided within a multimodal and intersemiotic message a concept linking "the mimetic and the diegetic level" (Moser 2007: 336) is needed and it is the concept of "iconicity that serves as a mediator" (*ibid*) between them.

Metaphor-icon link is always foregrounded in the communicated message, but its structure may be different depending on the prevailing component. The author of the research gives tribute to the model for metaphor-in-icon and icon-in-metaphor analysis in cognitive terms proposed by Hiraga in 1998, which is an effective tool "for clarifying the complex interrelationship between metaphor and iconicity in the dynamic production of meaning in language" (Hiraga 1998). To establish theoretical framework of the given research the works of Anstey and Bull (2006) on different types of semiotic systems, of Petrilli (2003) on intersemiotic translation, of Wagner (1996), Wolf (1999) and Moser (2007) on intermediality and multimodality, as well as of Forceville (1996) on pictorial metaphors shall be addressed.

The dynamically developing interdisciplinary field of environmental communication is a great platform for investigation of the role and application of the visual metaphor, as within the given field the tendency to apply visual metaphors for information transfer is quite widespread. Visual metaphors used in

different media to communicate information on environmental issues may frequently take a form of a social advertisement disseminated on posters, leaflets, and banners (both static and interactive). As a contribution to this idea, the author shall analyse the application of the visual metaphor '*global warming'* as appearing on both printed and web posters, created for the needs of public awareness campaigns. The paper sets out to explore the notion of a visual metaphor and its role in communicating specialised information to mass audiences, paying particular attention to the emotive and heuristic components of meaning construal. The paper shall focus on the investigation of simultaneous manifestation of iconic and metaphorical mappings, analysing metaphor-in-icon and icon-in-metaphor links in the visual metaphor '*global warming'*.

2. *Inter*-nature of Meaning: *Inter*disciplinarity, *inter*semioticity, *inter*mediality

Analysis of meaning creation processes within the field of environmental communication is complicated by the fact that the nature of information transferred has changed. There has been a considerable shift towards extensive *complexity* of meaning at all levels (formation, organization and transmission).

Successful application of visual metaphors for the needs of professional communication demands approaching the concept of *meaning* and its role in the communicative process from the threefold perspective – as an ***inter***disciplinary material, as an ***inter***semiotic construct, and/or as an ***inter***medial phenomenon. In other words, meaning within visual metaphors is efficiently decoded if a recipient is ready to address multiple scientific fields, recognise different symbols of multiple semiotic systems and accept information communicated via multiple media forms.

In the era when the borders between disciplines of knowledge are being blurred, addressing distinct scientific fields simultaneously is to a certain extent expected by the recipient of a message. The challenges occur due to the fact that this so called '*inter*-nature' of meaning creates a platform for communication of intersemiotic and increasingly intermedial and multimodal information. It demands a high level of multiliteracy on the part of the recipient, which can be seen on different planes and concerns the ability of the recipient to make a more efficient use of the general skills and professional competences to:

- manipulate imagery and visualize it, generating a bond between mental images and visual representation;
- think creatively, "using cognitive processes to manipulate or act on stored knowledge" (Nevid 2015: 251);
- establish certain relations to the previously acquired set of knowledge, i.e. to generate intertextual and intermedial links;

- to integrate perceptual and conceptual dimensions of meaning, where the former concerns the auditory, graphical and pictorial representation of the meaning-bearing component, e.g. word, compound, metaphor, etc.; while the latter refers to "the properties that constitute the meaning of the word" (following Massaro 1975: 12);
- demonstrate understanding of the isolated semiotic systems and of their combinations producing amalgam meaning, as "all semiotic resources are functional in the promotion of scientific learning" (Kress et al. 2001);
- deal with increasing (r)evolutionizing nature of content multimediality, which is seen as the gradual "passage from a received multimediality, to an interactive multimediality, to a built multimediality, to a reflective multimediality" (Dall'Aqua 2015: 219). In other words, it is a gradual transition from the content, which is pre-created, to the content, in the creation of which a user participates, to the content, which is created as a common social product, and, finally, to the self-creative content (cf. ibid).

Today, to create and communicate meaning in general and exploit visual metaphors in particular, a literate person "needs to understand both the conventions within each semiotic system and how combinations of semiotic systems work together to convey meaning within and across texts" (Marcuccilli Strop and Carlson 2010: 11). The signs used to produce and transfer meaning may belong to different semiotic systems which, according to Anstey and Bull (2006: 25), are the following:

> Linguistic (oral and written), visual (still and moving images), auditory (musical and sound effects), gestural (facial expression and body language), spatial (layout and organization of objects and space).

The author shares the opinion of Iyer and Luke (2010: 21), who treat "all forms of semiotic systems as texts". Moreover, the semiotic systems existing within one text "may overlap, co-occur, and work with or against the meanings of the other system(s)" (Albers and Murphy 2009: viii). It means that visual images can be read and recorded as texts, that images used to create visual metaphors "like texts are rhetorical and must use signs to express meaning" (Wagner 1996: 17).

Therefore, a multilitirate person should demonstrate a profound knowledge of semiotic theory, as "semiotics gives a foundation for the classification of signs according to different forms of iconicity" (May 2007 in Stathi 2014: 140). *Iconicity* or, as referred to by Wendorf (1990: 19), *iconicism*, implies, among all other factors, metaphorical incorporation of the elements and structures of one medium into the other (cf. ibid). This view is also supported by Stathi (2014: 147), who states that today within semiotic theories one should recognise that "…the making of meaning depends on, among other factors, processes of multimediality and intermediality…". In other words, in the communication of meaning encoded in the visual metaphor the elements of multiple semiotic systems are used

simultaneously creating a unified multimodal form and establishing intermedial links to the previously gained knowledge.

At this point, it is significant to review if the multiplicity of elements of different semiotic systems existing within one message and communicated via different media should be addressed as a *multimodal, multimedial, intermedial*, and/or simply *intersemiotic* phenomenon. The author of the paper is aware of the concepts of *trans-* and *plurimediality*, as well as *pluri-* and *transmodality*, but within the present paper she does not attempt to historicise all these concepts, but to propose operational definitions for the needs of the present research.

Multimodality, as defined by Kress and van Leeuwen (1996: 183), exists in "any text whose meanings are realised through more than one semiotic mode". The combination of various semiotic modes for communication of meaning implies application of different elements, different information representing resources, which are incorporated within one text. Therefore, the choice of different modes is justified not only by the necessity to pursue clarity of information, but rather is governed by the principles of essential "resource integration and meaning compression" (Baldry and Thibault 2006: 18-19, in Gibbons 2012: 17).

Multimediality, in its turn, has gained a considerable focus due to digitalization of information and is seen as one of the key characteristics of the communication on the Internet, alongside with hypertextuality and interactivity (cf. Ibáñez and Alonso 2015: 92). It refers to the "integration of different media formats, audio, graphics, pictures, and text, into a single media form" (Siapera 2012: 136), which is aimed at providing a better and faster access to the global knowledge, rather than to individual knowledge. According to Wolf (2005: 254), multimediality exists if "two or more media are overtly present in a given semiotic entity at least in one instance". It means that the existing media forms can still be clearly recognised and, if necessary, treated individually.

Intermediality "refers to interconnectedness of modern media of communication" (Klaus Bruhn Jensen cited in Donsbach 2008: 2385-2387). In other words, different media evident within a certain semiotic entity are mixed together to form a unified message, and, therefore, cannot be clearly delineated from each other, as they "depend on and refer to each other, both explicitly and implicitly" (*ibid*). Intermedial relations are significant for understanding and communicating the whole message, as they "integrate perceptual and conceptual dimensions of the construction of linguistic meaning" (Moser 2007: 342) and serve as "a bridge between medial differences that is founded on medial similarities" (Elleström 2010: 12). Therefore, the author of the present research following Wagner (1996) theories proposes to correlate the concept of '*intermediality'* with the concept of '*intertextuality'*, "where intertextuality is the mono-medial (verbal) and intermediality is the cross-medial variant" (Wolf 1999: 46), while both of them should be delineated as the forms of *intersemiotic relations*, existing between "different media in a multitracked text" (Nelson 2010: 13).

It means that irrespective of the fact whether transformation of the content from one form to another or simultaneous application of various media forms have been applied for creation of the message, it demands the recipient to understand the basic principles of transmutation or intersemiotic translation, as eventually it is all "a matter of perception and interpretation" (Nelson 2010: 13).

Discovering and interpreting intersemiotic messages makes a recipient experience a variety of different marginal emotions, e.g. joy-sadness, trust-disgust, fear-anger, anticipation-surprise, which can be further explored and analysed following, for instance, Plutchik's Theory of Emotions (1980). The emotional path to interpretation of the whole message and to the discoveries of new meanings is, perhaps, the fastest one, as most of our thoughts and actions are governed by emotions; but this path is not the most precise, as emotional response is always subjective and, frequently, hinders proper recognition of all semantic layers of meaning.

Instant and precise interpretation of visual metaphors is complicated by the fact that pictorial representation of the genuine object may to a certain extent be in conflict with recipients' perceptual and conceptual apparatus, i.e. it would appeal to emotions recipients are unwilling to experience, or would evoke allusions to the phenomena recipients are reluctant to consider. It means that the icon within a visual metaphor is seen as an efficient manipulative instrument, which is aimed at appealing directly to the aspirations, demands, beliefs and even immediate emotional state of the recipients, motivating them to act in a certain way, and, hence, influencing their behavior and decision-making. This assumption leads to the question, which shall be explored in the following sections, i.e.: How visual metaphors can be exploited to create a unique message, which communicates a universal challenge and is understandable irrespective of the linguistic background and professional expertise?

3. Visual metaphor

There is an ongoing discussion on the nature of visual metaphors, which are frequently seen as visual manifestations of corresponding verbal metaphors.

The author gives tribute to Forceville (1996) for the detailed critical survey of the studies addressing the concept of visual, or as termed by Forceville (*ibid*), pictorial metaphor, providing framework for its analysis and advocating in favour of Conceptual Metaphor Theory, which states that "metaphor is primarily a matter of thought and action, and only derivatively a matter of language" (Lakoff and Johnson 1980: 153). Theoretical and empirical research he conducted paved the way for future interdisciplinary investigations of the application of visual metaphors, which are seen as independent multimodal constructs that convey the message combining the elements of various semiotic systems.

Nevertheless, the sceptics still disregard the very notion of a visual metaphor, stating that "visual metaphors are nothing but the illustrations of commonplace

metaphors that already exist antecedently in language" (Caroll 2001: 202), i.e. they assert that there are no visual metaphors. This argument is supported by the fact that most of the visual metaphors can be relatively easily expressed on a verbal plane.

Visualisation of the verbal content establishing both metaphorical and iconic mappings is seen as being opposite to the verbalisation of the visual content, where iconic and metaphorical mappings are rather 'vividly explained' exploiting such rhetoric device as *ekphrasis*, which in general concerns "the verbal representation of visual representation" (Heffernan 1993). Reproduction of the image in verbal content is never precise, nor it is complete, as ekphrasis "is always less and more than the original visual sign" (Boehm 1995: 30). Eventually, ekphrasis is aimed at describing the visual artefact as vibrantly and precisely as possible, so that readers of this message can live the same emotions and feel the same aesthetic pleasure as if they saw the image themselves. However, it does not mean that verbal variant of the visually originated metaphor will retain metaphoric nature, as "description of metaphor simply does not have the power of the metaphor it describes" (Danto 1981: 172-173).

In other words, not all visual metaphors can be converted into verbal metaphors, and even if they can be reduced to the verbal statement, it does not guarantee the concordance of all its meanings in both semiotic systems.

> One mode's potential to render "meaning" can never be completely "translated" into that of another mode – and sometimes translation is downright impossible. For this reason alone, a healthy theory of (cognitive) metaphor must systematically study non-verbal and multimodal metaphor (Forceville and Urios-Aparisi 2009: 4).

Visualisation of a metaphor can make it broader or narrower than its verbal variant, it can establish a better rapport with the target audience due to clearly depicted meaning in use, finally, it can ensure a faster and better recognition of the metaphorically encoded meaning.

It means that a visual metaphor produces a unique story narrated through the mechanisms of both metaphorical and iconic mapping. Metaphorical and iconic mappings are simultaneously contributing to the transfer of a message, as the image used relates the pictorial dimension, linguistic meaning and "invites the user to consider its heuristic value" (Caroll 2001: 211). Users may identify new connections between the elements of the semiotic systems used to create a visual metaphor. They are invited to discover the stages of a visual metaphor construction, recognising "metaphoric mapping from the concrete to the abstract domain" (Demey et al. 2008: 194), and exploring "iconic mapping from the concrete source domain to the linguistic domain" (Demey et al. 2008: 194). It means that information encoded in the target domain is communicated metaphorically employing the "iconic depiction of the source domain" (Taub 2001: 97).

The mechanisms of iconic and metaphorical mapping are similar as they both "share that property of signification 'motivated' by similarity" (Hiraga 1998). However, similarity is both treated and manifested differently since in metaphorical mapping similarity is rather presupposed and/or represented in a schematised and abstract way. It is an ongoing never-complete process, as metaphorical links can be generated and/or established in the process of mapping one conceptual domain onto another conceptual domain (cf. Lakoff and Johnson 1980). In iconic mapping, similarity rather relies on object resemblance, replicating the shape, appearance or structure of an object, which makes icons "easy to understand because their resemblance to the objects is immediate and concrete" (Hiraga 1998).

> While "a metaphor, like an image or an analogy, is what it represents – but not because of an antecedent identity or similarity, not as a reminiscence, but in virtue of a similarity which it creates" (Anderson 1984: 459, in Hiraga 1998).

It means that metaphorical mapping creates almost infinite number of meanings and their possible interpretations, as "the particular *content* of a metaphor can be said to constitute an interpretation of reality in terms of mental icons that literally allows us to *see* what is being talked about (italics in the original)" (Danesi 1995: 266). Moreover, metaphorical mappings can even lead to interpretations, which were not intended by the authors of the message.

In visual metaphors, iconic representation is frequently seen as being dominant over the metaphorical component in the overall meaning construction, as the links established between the icon and the object it resembles are evident to people irrespective of their background knowledge and expertise. However, "it is a mistake to suppose that visual metaphors can only induce object-comparison and cannot mobilize knowledge of connotations" (Caroll 2001: 359). The pleasure and great aesthetic value in creating and using visual metaphors lie in the fact that having initiated user imagination and evoked so many ideas and associations simultaneously, these metaphors nevertheless are not "being reduced to any definite thought" (cf. Kant 1790 (1964): 317-318). This factor makes visual metaphors a universal tool for attracting attention and communicating information irrespective of the linguistic competence, professional expertise and background knowledge of the recipient, as these metaphors are applied to address the shared challenges and they are constructed manipulating commonly recognised images.

Within the field of environmental communication, the tendency to apply visual metaphors for information transfer is quite widespread, due to explicit iconicity of the very concept '*environment'* and related concepts of '*nature'* and '*life'*. The advocates of 'green thinking' tend to visualize information on environmental issues, which may frequently take a form of a social advertisement being disseminated on posters, leaflets, and banners within public awareness campaigns. This phenomenon can be well illustrated with the widespread use of the visual metaphor '*global warming'*.

4. Case Study: Metaphor-Icon Links in Decoding Visual Metaphor 'Global Warming'

Metaphor-icon links within visual metaphors can be based on metaphor-in-icon and icon-in-metaphor models (following Hiraga, 1998). The icon-in-metaphor model assigns dominant role to iconic mapping, which demands addressing the visual representation, i.e. the icon to construct the meaning of the whole metaphor. In this case iconicity encompasses the metaphorical component and metaphor is manifested on the visual plane, as clearly as no multiple readings are possible.

While the metaphor-in-icon model is a more complicated mechanism, as it implies primarily addressing the verbal element, first to relate it to iconic representation and, secondly, to grasp the meaning of the whole visual metaphor. The difficulty lies in the fact that decoding verbal message may lead to multiple interpretations and a user should be able to link the intended meaning of the verbal component with the provided visual element to properly understand the communicated message.

However, despite the fact that a user of the message is manipulated to think of a certain idea, s/he is not guided to experience only a restricted set of emotions, and so s/he is not controlled over the discoveries s/he can make while reading the message. This assumption increases the role of emotive and heuristic components in creating visual metaphors in meaning formation and communication processes, which shall further be addressed analysing particular examples of a visual metaphor '*global warming'* in use.

The '*global warming*' metaphor is iconic by its very nature, as it is best comprehended in the pictorial dimension, because it creates a visual string of related images, which lead to complete understanding of the meaning encoded within the metaphor. Explicitness of the icon grants a deliberate freedom to the users of this metaphor to visualize its meaning components exploiting various semiotic systems and combining different communicative modes. It leads to generally numerous variants of a visual metaphor '*global warming*', which, although, definitely share the key components and are aimed at achieving a unified result, are still communicated employing different means and appealing to different emotions.

The image of the Earth in the visual metaphor '*global warming'* is frequently used in combination with signs belonging to other *symbolic* systems or with elements belonging to other *semiotic* systems, thus, producing a unique story narrated through the mechanisms of metaphorical and iconic mapping. These actions tend to evoke a wide range of emotions, which, however, all should lead to the idea of negative, even tragic effects global warming produces on the Earth. The negative effects produced on the planet are visualized as the Earth *melting, sweating, heating, burning, boiling, warming, crying, being in a fever*, in other words, *secreting* its essential fluids.

The image of the planet Earth in the visual metaphor '*global warming*' may be represented in different ways and forms, e.g.:

- *photo* – the genuine photo of the planet Earth accompanied with some additional symbols and/or verbal message;
- *cartoon* – an illustration of a personage with characteristics of a human being;
- *graphical image* – the schematic representation of the planet made using various graphical editors (usually implies combining elements of various systems);
- *painting* – an authorised artistic representation of an image of the Earth in pastel or oil colours (usually produced to pronounce the attitude of an artist towards a particular problem);
- *drawing* – an artistic creation of an image of the Earth (usually authorised and produced for the needs of a particular event).

The best effect is achieved when the authors of the posters choose to adjust the real photos of the planet Earth to the needs of their public awareness campaigns. The photos of the planet may be enhanced, artistically adjusted, their quality may range from blurred to superior, but the meaning they convey is still clearly understandable. Such pictures appeal directly to our knowledge of how the planet looks like and trigger the strongest emotions, as they are created as if depicting the current situation. Within the framework of the present research, the author focuses on the analysis of different posters containing iconic images of the planet Earth. To interpret the meanings communicated by the images on posters, advanced level of multiliteracy is expected on the part of the recipient, which requires the ability to think creatively, manipulate images, visualize associations, and generate intertextual and intermedial links.

In Figure 1, the Earth is shown as a melting ice cream. The image is a typical representation of the icon-in-metaphor model, as for the interpretation of the whole text it is most relevant to understand the visual code placed in the spatial setting. The planet Earth as the image of an ice cream placed against the dark surrounding background with the enlightened middle part may be interpreted as the light of still existing hope for the better life, if the proper actions are taken as soon as possible.

Figure 1. The Earth as the melting ice cream[1]

The same message can be communicated employing the metaphor-in-icon model, which would definitely demand addressing the verbal message first, assess it against the visual and spatial settings and, by adding these dimensions, interpret it.

Figure 2. The Earth as the melting ice cream, text added[2]

The image given in Figure 2 is a great example of the intersemiotic message communicated on linguistic, visual and spatial planes. The reader is invited to notice the slogan, which is a homophonic pun of "*I Scream*" vs. "*Ice cream*". The latter is supported visually to guide the reader to the required interpretation. The image of the Earth is enhanced and artistically treated to intensify the effect if compared against the picture given in Figure 1, while the image of the waffle cone remains unadjusted. The message "*not enough*" written in small letters just beneath the slogan indicates insufficiency of the actions taken so far to fight the problem. It is interesting to note that in Figure 2 the visual metaphor is actually

[1] Available from: http://vivredemain.fr/35-campagnes-creatives-wwf-reflechir. [Accessed: 30 May 2017].

[2] Available from: http://www.hongkiat.com/blog/global-warming-alert-posters. [Accessed: 30 May 2017]

named, as the slogan "*global warming*" is applied at the top of the poster, which, however, is not needed as the communicated message is understandable without tautological repetition of the same idea.
In Figure 3, the planet is frying on the pan, gradually losing its shape and essential fluids in the form of vapour. In the given case, the visual metaphor is constructed employing the elements of visual, acoustic and spatial semiotic systems.

Figure 3. The Earth on the frying pan[3]

If compared to the images communicated in Figures 1 and 2, the image seems to be less powerful and rather static, as time component is not well pronounced. This image created following the icon-in-metaphor model lacks the semantic component of inevitability, as it completely depends on the humanity whether to fire a pan or not, while in the case of ice cream, the balance is created artificially and is very fragile, as the ice cream is supposed to melt and disappear (be eaten) anyway irrespective of the fact if it is or it is not treated properly.

The use of icon-in-metaphor model for visual metaphor creation may imply application of the elements belonging to distinct symbolic media. This phenomenon is illustrated in Figure 4, where the photo of the planet Earth is hidden behind the transparent traffic sign NO ENTRY. The visual metaphor communicates multiple messages simultaneously and demands the user to understand *intermedial* nature of the visual metaphor to uncover its meaning(s). On the one hand, the use of the traffic sign clearly implies reasonable limitations on the use of vehicles, which produce emissions of CO_2 to the atmosphere, hence, contributing to intensifying global warming effects. On the other hand, the traffic sign restricts the use of any vehicles including environment-friendly bicycles, which can be treated as an invitation to reconsider the very concept of green logistics, or even as a warning not to welcome any humans. Placement of the traffic sign over the image of the whole planet indicates the scale of the problem.

[3] Availbale from: https://www.123rf.com/photo_6587701_global-warming-concept-earth-at-the-frying-pan-isolated-on-white.html. [Accessed: 30 May 2017]

The message is intensified by the choice of colours for the slogan and background. The colours of the slogan "*global war**n**ing*", which is seen as a pun to "*global war**m**ing*", are the same as the colours of the road sign, which are placed against the very dense black background. The red colour for "war***n***ing" is used on purpose to allude to "war***m***ing", intensifying the meaning of being very warm, even hot.

Figure 4. The Earth bearing NO ENTRY sign[4]

The combination of different visual and verbal factors intensifies the meaning of the multimodal message and stimulates the thinking process, inviting both intertextual and intermedial links, as well as welcoming unique and expected interpretations. The use of verbal components on the poster makes the user consider either words or image(s) in the first order, which would influence the interpretation path.

Figure 5 illustrates the abovementioned phenomenon. A human hand holding the lighter, which heats the planet illustrates the influence of the mankind on the Earth. The question "Why?" put at the bottom of the poster is a rhetoric one. The slogans placed on the posters may frequently take the form of the rhetoric open questions with multiple subjective answers. The questions pronounced on the posters suggest acoustic perspective, as depending on the way you literally ask it you may induce different reaction within the audience.

[4] Available from: https://uk.pinterest.com/pin/553379872935921646/. [Accessed 14 June 2017]

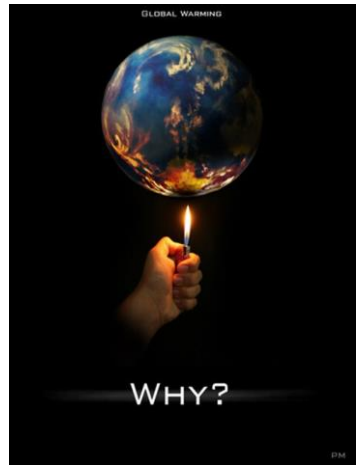


Figure 5. The Earth being heated by a lighter [5]

The concept of fire as the summative visual representation of the effect produced by human beings on the planet Earth is applied in versatile forms. The use of the image of fire is unambiguous allusion to the concept of *global warming*, which is clearly associated with reaching high temperatures.

Figure 6 contains the image of the Earth, which is represented as the melting candle. The visual metaphor in Figure 6 is created following the metaphor-in-icon model, with the allusion to the idiomatic expression "Plan B", which stands for the backup plan in case plan A fails. In the given case, the visual metaphor of the melting candle is applied to intensify the meaning expressed on the linguistic plane, as the role of the verbal message is dominant. The slogan goes "*There is no planet B* [intermission, no syntactic signs] *Act now*.", which invites the target audience to consider the tragic consequences of their inactivity and even careless laziness. The component of time is pronounced in both visual and verbal messages. In the former, the influence of time is illustrated with the burning candle, as the time required for the candle to burn down completely is very short. In the verbal message, the semantic component of time is indicated with the full stop at the end of the sentence "*Act now*.", specifying the absence of time for decision-making. This dual indication makes the sense of inevitability of the tragedy evident. However, the very image communicates the illusion of the existing control over the situation, which is supported by the fact that we can decide whether to light up the candle.

[5] Available from: https://www.pinterest.com/pin/518336238339889740/. [Accessed 16 June 2017]

Figure 6. The Earth as the burning candle[6]

It is interesting to note that both Figure 5 and Figure 6 are created employing the image of open fire against the dense black background. However, in Figure 5 black background colour is required to intensify the damage caused by the fire of the lighter (as flame is better visible against dark background), while in Figure 6 the flame of the candle creates a path of light and points at the slogan and the exact date, inviting people to start thinking *green* first on the Earth Day on April 22, and later continue maintaining the same attitude throughout their lives.

The call for certain action, which in the opinion of public awareness campaign authors should be taken by the humans, can be expressed in the form of directives, which invite and even command people to act in a certain way. The verbal message is then supported by the visual message, establishing a clear metaphor-in-icon link between them.

The verbal metaphor used in Figure 7 "*Let the Earth sleep*" can be interpreted as an invitation, an order, a command, an advice, a request; finally, as a plea ranging from the woeful cry to horrid yell, which requires addressing the acoustic perspective for the interpretation of the multimodal message. The concept of *sleeping* reflected in the verbal metaphor conceptually suggests treating the Earth as the human being, and implies such components as proper rest, absence of acoustic and visual pollution, and full restoration, which are essential to ensure sound living.

---

6 Available from: https://www.pinterest.com/pin/562668547176596438/. [Accessed on 16 June 2017]

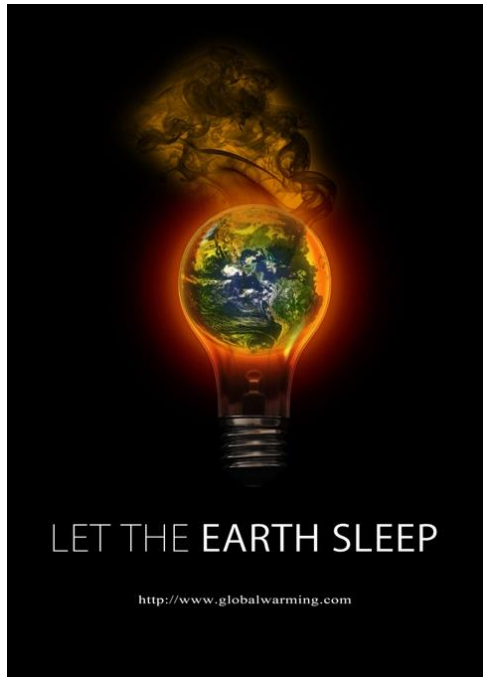


Figure 7. The Earth as the burning light bulb [7]

The visual metaphor of *sleeping* is expressed with the help of the image of a light bulb. The metaphor has a dual reading, as on the one hand, the light bulb represents the surrounding acoustic, visual and other kinds of pollution, which disturb the Earth from having rest. For the proper rest, the light bulb should be switched off, thus creating the desired atmosphere of pleasant silence and nightfall. While, on the other hand, the light bulb can be treated as the symbol of the planet itself, which is heated to the utmost and emits black and flame-colour vapours, which most probably should be accompanied with unpleasant sound of light bulb breaking (thus adding the acoustic dimension), as well as with unpleasant smell of light bulb burning. In this case, the light bulb placed against the very dense black background, is used, in a way, to symbolise the powerlessness and illustrate the inability of the light bulb to enlighten the surrounding. It can be read as the absence of hope if the Earth is not allowed to revive.

Some of the posters used for the needs of social campaigns do not actually invite to take actions, but simply state the current situation (see Figure 8). The power of such visual metaphors is hidden in the fact that they simultaneously allow minimum interpretive freedom (containing clearly expressed idea) being

[7] Available from: http://www.coroflot.com/capioxy/Global-Warming-Poster-Graphic-Design. [Accessed on 15 June 2017]

highly subjective and evoking the emotions, which many people are reluctant to think of and experience.

Figure 8. The Earth on fire and secreting the fluids[8]

In Figure 8, a human is holding the planet Earth in one's hands, the planet is burning on the top and is secreting the essential fluids at the bottom. There is a visual metaphor of *the Earth in a way bleeding* hidden within the visual metaphor of global warming. Just as the human is committed to die having lost a substantial amount of blood, the Earth is going to perish having lost its water resources. The slogan used for the present poster states "*We did this.".* The full stop at the end of the slogan is very unusual and in this case it is used to escape any other readings and interpretations of the given statement, which can be interpreted as "***This is it***. 'We did this***. No one else to blame***". The sense of tragedy present in the given figure leaves very little for the light of hope. The poster is basically the last and final call to take the required actions, which were implicitly pronounced and hinted in Figures 1 – 3, and were obviously invited and proposed in Figures 4 – 7, analysed within the framework of the present paper.

The decision to visualize environmental issues on printed and web posters is substantiated by the fact that information dissemination capability of the given medium is very high. The '*global warming*' issue is reflected in numerous artistic works, papers, posters, advertisements, and is supported via an impressive number of poster competitions and advertisement festivals.

[8] Available from: https://www.behance.net/jacobbrcic. [Accessed 20 June 2017]

5. Concluding Remarks

The intersemiotic approach to message creation shifts our perception of data collection, procession, and representation, welcoming active use of the elements belonging to distinct symbolic media and different semiotic systems, as a result producing texts with the presence of the elements of *linguistic*, *visual*, and *spatial* systems, and even with the imaginative manifestation of the elements of the *acoustic* system (implying considering the tone, pauses, pace and intonation when reading the message) and *gestural* system (considering the expressions and gestures in motion). Visual metaphors are generally constructed simultaneously encompassing the elements of at least three semiotic systems, hence creating a unified multimodal form and establishing intermedial links to the previously gained knowledge.

Application of visual metaphors, created following both metaphor-in-icon and icon-in-metaphor models, for communication of sensitive issues within the framework of public awareness campaigns is a very powerful tool, which can ensure equal effect in different linguistic communities, producing a unique story narrated through the mechanisms of metaphorical and iconic mapping. The use of the icon-in-metaphor model is frequently given priority due to its less ambiguous nature, as the links established between the icon and the object it resembles are transparent to people irrespective of their mother tongue, expertise, and to a certain extent, irrespective of their background knowledge.

Visualisation of '*global warming'* effects raises topical issues in an understandable, concise and very precise way, ensuring maximum information coverage framed by minimum interpretive guidelines. Therefore, interpretation of the text communicated by visual metaphor is a fascinating exercise, as despite the fact that it seems to express the obvious reality, such text encompasses unique heuristic value and aesthetic pleasure in discovering multiple layers of meaning.

References

Albers, Peggy and Sharon Murphy. 1999. *Telling Pieces. Art as Literacy in Middle School Classes*. Taylor and Francis.

Anderson, Douglas. 1984. Peirce on Metaphor. *Transactions of the Charles S. Peirce Society: A Quarterly Journal in American Philosophy 20(4)*, 453-468.

Anstey, Michelle and Geoff Bull. 2006. *Teaching and Learning Multiliteracies: Changing Times, Changing Literacies*. Newark, DE: International Reading Association.

Boehm, Gottfried. 1995. Bildbeschreibung. Über die Grenzen von Bild und Sprache. In Boehm, Gottfried; Pfotenhauer, H. (Hg.): Beschreibungskunst - Kunstbeschreibung. Die Ekphrasis von der Antike bis zur Gegenwart, München, 23-40.

Caroll, Noel. 2001. *Beyond Aesthetics*. Philosophical Essays. Cambridge University Press.

Cox, Robert, and Phaedra C. Pezzullo. 2015. *Environmental Communication and the Public Sphere,* 440. SAGE Publications; 4th ed.

Dall'Aqua, Luisa. 2015. Orientism Management (OM): A New Framework to Manage Decisions and Hyper Dynamic Knowledge Process in a MultiUser Network. In Alok Kumar Goel and Puja Singhal (eds.), *Product Innovation through Knowledge Management and Social Media Strategies.* IGI Global.

Danesi, Marcel. 1995. The Iconicity of Metaphor. In Marge E. Landesberg (ed.), *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*, 265-284. Berlin: Mouton de Gruyter.

Danto, Arthur. 1981. *The Transfiguration of the Commonplace*. Cambridge, MA: Harvard University Press.

Demey, Eline, Van Herreweghe Mieke and Myriam Vermeerbergen. 2008. Iconicity in Sign Languages. In Klaas Willems and Ludovic De Cuypere (eds.), *Naturalness and Iconicity in Language*, 189-214. John Benjamins Publishing Company.

Elleström, Lars. 2010. The Modalities of Media: A Model for Understanding Intermedial Relations. In Lars Elleström (ed.), *Media Borders, Multimodality and Intermediality* 11-48. Hound- mills, Basingstoke and Hampshire: Palgrave MacMillan.

Forceville, Charles. 1996. *Pictorial Metaphor in Advertising.* Routledge.

Forceville, Charles and Eduardo Urios-Aparisi (eds.). 2009. *Multimodal Metaphor*. Applications in Cognitive Linguistics, Vol. 11. Berlin: Mouton de Gruyter.

Gibbons, Alison. 2012. *Multimodality, Cognition and Experimental Literature*. Routledge studies in Multimodality: 3. Routledge.

Heffernan, James A. W. 1993. *Museum of Words: The Poetics of Ekphrasis from Homer to Ashbery*. Chicago: University of Chicago Press.

Hiraga, Masako K. 1998. Metaphor-icon Link in Poetic Texts: A Cognitive Approach to Iconicity. *Journal of the University of the Air*, No. 16. 95-123.

Ibáñez, Daniel Barredo and Martin Oller Alonso. 2015. Multimedia resources and Sporting Bias in MARCA.com and SPORT.es: An Analysis of Matches between Real Madrid C.F. and F. Barcelona in 2010-11. In Simon Gwyn Roberts (ed.), *Sport, Media and Regional Identity*, 91-104. Newcastle upon Tyne: Cambridge Scholar Publishing.

Iyer, Radha and Luke Carmen. 2010. Multimodal, Multiliteracies: Texts and Lietracies for the 21st Century. In Darren L. Pullen and David R. Cole (eds.), *Multilietracies and technology Enhanced Education: Social Practice and the Global Classroom,* 18-34. IGI Global.

Jensen, Klaus Bruhn. 2008. Intermediality. In W. Donsbach (ed.), *International Encyclopedia of Communication*, 2385-2387. Malden, MA, USA: Blackwell Publishing Ltd.

Kant, Immanuel. 1790. [1964]. The Critique of Judgement. In Hofstadter and Kuhns (eds.), *Philosophies of Art and Beauty: Selected Readings in Aesthetics from Plato to Hedegger*, New York: Modern Library.

Kress, Gunther and Theo van Leeuwen. 1996. *Reading Images: The Grammar of Visual Design*, London: Routledge.

Kress, Gunther, Jewitt, Carey, Ogborn, Jon and Charalampos Tsatsarelis. 2001. *Multimodal Teaching and Learning: The Rhetorics of the Science Classroom*. London: Continuum.

Lakoff, George and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: Chicago University Press.

Marcuccilli Strop, Janice and Jennifer Carlson. 2010. *Multimedia Text Sets. Changing the Shape of Engagement and Learning*. Portage and Main Press.

Massaro, Dominic W. 1975. *Understanding Language. An Information-Processing Analysis of Speech Perception, Reading and Psycholinguistics*. Academic Press, INC.

May, Michael. 2007. A Semiotic Framework for the Semantics of Digital Multimedia Learning Objects. 14th International Conference on Image Analysis and Processing Workshops, 2007. ICIAPW 2007. 33-38. Los Alamitos, CA: IEEE Computer Society Press.

Moser, Sybylle. 2007. Iconicity in Multimedia performance: Laurie Anderson's White Lily. In Elżbieta Tabakowska et al. (eds.), *Insistent Images*, 232-347. John Benjamins Publishing Company.

MWD – Merriam Webster Dictionary – Available from: http://www.merriam-webster.com/dictionary/communication. [Accessed on 9 June, 2017].

Nelson, Robin. 2010. Introduction: Prospective Mapping and Network of Terms. In S. Bay-Cheng, C. Kattenbelt, A. Lavender and R. Nelson (eds.), *Mapping Intermediality in Performance*, 13-24. Amsterdam University Press.

Nevid, Jeffrey, S. 2015. *Essentials of Psychology. Concepts and Applications*. Cengage Learning.

Petrilli, Susan. 2003. *Translation, Translation*. Amsterdam and New York, Rodopi.

Platonova, Marina. 2015. Applying Emotive Rhetorical Strategy to Environmental Communication in English and Latvian. In *Procedia: Social and Behavioural Sciences*, 107-113. Vol. 236., Elsevier.

Plutchik, Robert. 1980. *Emotion: Theory, Research, and Experience*: Vol. 1. Theories of Emotion 1. New York: Academic Press.

Purchase, Helen and Daniel Naumann. 2001. A Semiotic Model of Multimedia: Theory and Evaluation. In Syed Mahbubur Rahman (ed.), *Design and Management of Multimedia Information Systems: Opportunities and Challenges*, 1-21. Idea Group Publishing.

Siapera, Eugenia. 2012. *Understanding New Media*. SAGE Publications LTD.

Stathi, Irini. 2014. Film and New Art Media Semiotics: On the Figural. In Evripides Zantides (ed.), *Semiotics and Visual Communication: Concepts and Practices,* 139-149. Newcastle upon Tyne: Cambridge Scholar Publishing.

Taub, Sarah. 2001. *Language from the Body. Iconicity and Metaphor in American Sign Language*. Cambridge: CUP.

Wagner, Peter. 1996. Introduction: Ekphrasis, Iconotexts and Intermediality – the State(s) of the Art(s). In Wagner Peter (ed.), *Icons-Texts-Iconotexts. Essays on Ekphrasis and Intermediality*, 1-42. Walter De Gruyter & Co.

Wendorf, Richard. 1990. *The Elements of Life: Biography and Portrait-painting in Stuart and Georgian England.* Clarendon Press.

Wolf, Werner. 1999. The Musicalization of Fiction: A Study in the Theory and History of Intermediality. Internationale Forschungen zur Allgemeinen und Vergleichenden Literaturwissenschft 35. Amsterdam: Rodopi.

Wolf, Werner. 2005. Intermediality. In David Herman, Manfred Jahn, and Marie-Laure Ryan (eds.), *Routledge Encyclopedia of Narrative Theory,* 252-256. London: Routledge.

PROSODICALLY-CONDITIONED SYLLABLE STRUCTURE IN ENGLISH*

PAULA ORZECHOWSKA
Adam Mickiewicz University, Poland
paulao@wa.amu.edu.pl

JANINA MOŁCZANOW
University of Warsaw, Poland
jmolczanow@uw.edu.pl

MICHAŁ JANKOWSKI
Adam Mickiewicz University, Poland
mjank@wa.amu.edu.pl

Abstract
This paper investigates the interplay between the metrical structure and phonotactic complexity in English, a language with lexical stress and an elaborate inventory of consonant clusters. The analysis of a dictionary- and corpus-based list of polysyllabic words leads to two major observations. First, there is a tendency for onsetful syllables to attract stress, and for onsetless syllables to repel it. Second, the stressed syllable embraces a greater array of consonant clusters than unstressed syllables. Moreover, the farther form the main stress, the less likely the unstressed syllable is to contain a complex onset. This finding indicates that the ability of a position to license complex onsets is related to its distance from the prosodic head.

Keywords: English, lexical stress, onsets, phonotactics

1. Introduction

Traditionally, syllable weight has been associated with rhyme. In *The Sound Pattern of English*, Chomsky and Halle (1968) stated that stress is attracted to heavy syllables, defined by means of a rhyme structure, where onset is considered not to affect stress placement. This observation was confirmed in a large body of cross-linguistic research (e.g. Clements and Keyser 1983; Halle and Vergnaud 1987; Hyman 1985; Hayes 1995; Selkirk 1984). However, the studies that followed reported on systems in which not only the presence of onset but also its quality determine the position of stress. For instance, in Western Aranda, an

* This research has been financed by the National Science Centre, Poland, under grant no. 2015/18/E/HS2/00066 titled *The role of phonological features in phonotactics: A study on structure and learnability of consonant clusters in Slavic and Germanic languages*.

Australian language, words composed of more than two syllables are assigned initial stress if they start with a consonant (Davis 1988). A similar pattern was found in Pirahã, an Amazonian language, where the voice specification of onset consonants affects the stress placement (Everett and Everett 1984). Recent comprehensive overviews of such systems additionally list Karo, Banawã, and Dutch (see Topintzi 2010 for references and discussion).

Although English has been generally mentioned among languages which exhibit no interaction of stress with the onset quality or the presence of onset, recent studies (Kelly 2004; Ryan 2014) have revealed that onset complexity (i.e. length) is a significant attractor of stress in English bisyllabics. In the present study, we examine the correlation between stress placement and the structure of word-initial onsets in polysyllabic words with the main stress located on different syllables (word-initial, word-medial and word-final). Note that complexity is understood here in terms of the number of adjacent consonants, and it ranges from a zero onset (onsetless), through a simplex onset (represented by a single consonant, C) to a complex onset (represented by a cluster of consonants, CC and CCC).

2. State of the art

Kelly (2004) reported on three studies investigating the relationship between word-initial onset length and stress pattern. Firstly, the analysis of a corpus-based list of 6 862 items (including 4 126 trochaic and 2 736 iambic words) demonstrated that the first syllable (or trochaic) stress propensity is significantly related to the number of consonants word-initially. The tendency is expressed by means of trochaic proportions, which ascend from 0.35 through 0.69 and 0.83 to 0.98 for zero-, one-, two- and three-member onsets, respectively. The same tendency was observed in psycholinguistic data and in poetry. Kelly (2004) showed that in a judgement test English speakers relied on word onset structure in determining the stress of nonce bisyllabics, thus suggesting that the association between onset complexity and stress placement is part of native speakers' phonological representation. Finally, onset structure was shown to affect the alignment of words with stressed positions in John Milton's *Paradise Lost.* Generally, these findings indicate that onset patterns can have wide ranging effects on English prosody.

Building on the contribution of Kelly (2004), Ryan (2014) analysed a larger set of morphologically simplex English bisyllabics (N = 8 323), extracted from CELEX (Baayen et al. 1993), controlling for variables such as part of speech (noun, adjective, verb), frequency range (low, middle, high), onset structure (zero, C) and rhyme structure (a short vowel V, CV, a long vowel or a diphthong VV). Overall, there is a significant correlation between stress placement and word onset structure, which persists across all the major subdivisions of the lexicon. A

summary of the results is presented in Figure 1. The percentages refer to bisyllabic words stressed on the initial syllable with initial onset length varying from zero onset (labelled 0) to a three-member consonant cluster (labelled 3).

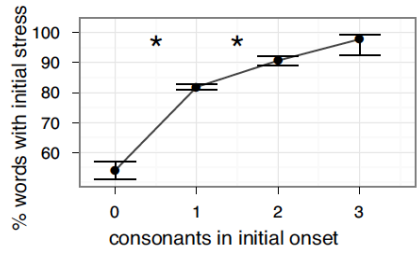


Figure 1: Stress placement in English bisyllabics of varying onset structure (Ryan 2014: 311)

The ascending line confirms that the incidence of initial stress increases monotonically with an increase in onset complexity. That is, the more complex the onset, the more likely the syllable is to carry primary stress. Each asterisk in the plot indicates a significance level of a pairwise comparison between the proportions below it (with 95% confidence intervals). As can be observed, the differences between onset zero and C, as well as between C and CC are significant. With these results in mind, we pursue an analysis which looks into the correlations between the onset type and stress pattern on a more exhaustive and varied dataset.

3. Present study

The present study examines the relationship between stress patterns and onset complexity. Emphasis is placed on onset clusters found in the word-initial syllable only. Target words collected for the purpose of the study vary in length from one to eight syllables. Such a rich and varied inventory of words makes it possible to investigate the relation not only in a binary contrast between stressed and unstressed syllables (cf. Kelly 2004; Ryan 2014), but also accounting for the degrees of distance from the main stress, in line with Orzechowska et al. (2018). We are particularly interested in finding out whether the complexity of the initial onset is affected by its distance from the main stress. The following labelling convention has been adopted, where the initial syllable is underlined (_), and the primary stress is marked with an apostrophe (’):

(1) a. *bisyllabics*
'al.mond – 0 distance between onset and stress
re. 'verse – 1 distance
b. *trisyllabics*
'bi.tter.ness – 0 distance
em. 'ploy.ment – 1 distance
pro.me. 'nade – 2 distances
c. *tetrasyllabics*
'bu.dge.ta.ry – 0 distance
e. 'mo.tion.less – 1 distance
co.ro. 'na.tion – 2 distances
te.le.gra. 'phese – 3 distances

3.1. Data

The data were extracted from two resources, a dictionary and corpus. The list of unique words was extracted from the 85 000-entry *Oxford Advanced Learner's Dictionary of Current English* (henceforth ALDCE, Hornby 1974). We eliminated proper nouns, abbreviations and duplicated entries, which resulted in the total of 26 614 uninflected lexical entries stressed on the first syllable, and starting either with a vowel (onsetless), a simplex onset (C) or a complex onset (CC, CCC). The data included compounds and morphologically complex words belonging to different parts of speech. The final word list contained 26 614 lexical items, including 3 730 monosyllabics, and 22 884 polysyllabics.

The dictionary data provided us with word type frequencies, where each lexical item had the frequency equal to one. In order to complement the analysis with a richer dataset, token frequencies for the word list were extracted from the 410 million-word *Corpus of Contemporary American English* (henceforth COCA, Davies 2011). This corpus of newspaper texts contained words and their repetitions, reaching a total of 239 291 747 word tokens, among which 159 467 346 monosyllabics and 79 824 401 polysyllabics were found.

3.2. Results

Below, we summarize the results of the analysis for words of various lengths composed of up to eight syllables. Sections 3.2.1 to 3.2.5 report on the ALDCE data where word type frequencies are given, while section 3.2.6 generalizes over the corpus-based token frequencies. In each table, the Onset (O) column specifies an onset type of the word-initial syllable, which can have the following structure:

(2) a. onsetless (labelled 0)
b. simplex (containing a singleton consonant; labelled 1)
c. complex (containing a sequence of consonants CC, CCC and CCCC; labelled 2, 3 and 4, respectively).

The distance column (Dist) presents the number of words with a specific distance between the stressed syllable and the first syllable, ranging from one to five. Distance zero (Dist 0) represents words stressed on the first syllable (e.g. *'tro.lley, 'ar.bi.tra.ri.ness*), while distances one to three (Dist 1-3) embrace words with the main stress on the second, third and fourth syllables (e.g. *beau. 'ti.cian, in.di. 'vi.dual, im.po.ssi. 'bi.li.ty*), respectively (see section 3). Tables 1-5 report on the number of words (No), and percentages (%) calculated for the total of words with a particular stress pattern (Σ). For instance, in bisyllabics (see Table 2) with trochaic stress ['..] (Dist 0), 9% of words have no onset, 69% start with C, 21% start with CC, and only 1% of words contain a CCC-initial cluster. In bisyllabics with iambic stress [.'.] (Dist 1), 36%, 56% and 8% of words contain zero, one or two consonants, respectively. Only two lexical items with iambic stress start with a three-member cluster.

In order to verify whether the differences observed in Tables 1-5 hold, a statistical analysis was performed. The Pearson's chi-squared test (χ^2) demonstrates that the differences between onset length and stress placement are statistically significant at a high level ($p < .000$). It must be noted that χ^2 does not operate on low numbers. Therefore, only categories represented by more than five words were entered into the analysis.

3.2.1. Monosyllabics

The discussion of the results starts with words composed of a single syllable, all of which are characterized by a distance of zero (Dist 0). In other words, all monosyllabics have primary stress. Therefore, Table 1 shows the distribution of various onset types in stressed syllables, and constitutes the baseline for the comparison with longer words.

Table 1: Monosyllabic words in the dictionary (N = 3 730)

	Dist 0	
Onset	**No**	**%**
0	129	3
1	2 338	63
2	1 154	31
3	109	3
Σ	3 730	100

As can be observed, the majority of monosyllabics starts with a single consonant (63%). The number of words gradually decreases from 63% to 3% along with the increase in onset complexity from C to CCC. Onsetless initial syllables constitute a minority, and are found only in 3% of the monosyllabic words under investigation.

3.2.2. Bisyllabics

In bisyllabics, the unstressed position (Dist 1) contains fewer words with an initial consonant cluster than the stressed one (Dist 0). That is, more words start with a sequence of two or three adjacent consonants when the initial syllable is stressed ($p < .001$). As is shown in Table 2, the proportion of CC(C) in a stressed syllable in relation to an unstressed syllable amounts to 11:1 (1 742 to 161) for CC and 60:1 (119 to 2) for CCC. This relation can be also expressed by trochaic proportions (labelled TP), which increase with the number of consonants in the word onset.

Table 2 Bisyllabic words in the dictionary ($N = 10\ 493$) with initial main stress as a function of initial onset complexity

	Dist 0		**Dist 1**		
Onset	**No**	**%**	**No**	**%**	**TP**
0	730	9	742	36	.49
1	5 850	69	1 147	56	.83
2	1 742	21	161	8	.91
3	119	1	2	0	.98
Σ	8 441	100	2 052	100	

The results are in accordance with the previous findings in Kelly (2004) and Ryan (2014), where onset complexity was demonstrated to attract stress in English bisyllabics. It must be noted that the propensity of cluster-initial words to exhibit a trochaic pattern is greater in shorter words. Therefore, in order to establish whether the same tendencies hold over the entire lexicon in English, below we detail the analysis for polysyllabic words.

3.2.3. Trisyllabics

Similarly to trochaic bisyllabics, trisyllabics accented on the first syllable feature more complex onsets ($p < .001$). In addition, the data in Table 3 show that the distance is correlated with the degrees of phonotactic complexity; greater proximity to the stressed syllable guarantees greater tolerance of consonant clusters. At the same time, the farther away from the main stress, the greater the number of onsetless syllables. The same tendency, although less pronounced, was observed in the group of bisyllabics.

Table 3: Trisyllabic words in the dictionary (N = 7 126) with initial main stress as a function of initial onset complexity

	Dist 0		Dist 1		Dist 2	
Onset	**No**	**%**	**No**	**%**	**No**	**%**
0	736	19	961	37	304	48
1	2 522	66	1 400	53	298	47
2	559	14	273	10	35	5
3	29	1	8	0	1	0
Σ	3 846	100	2 642	100	638	100

3.2.4. Tetrasyllabics

The results obtained for tetrasyllabics strengthen the findings reported in sections 3.2.2 and 3.2.3. The number of words with an initial CC cluster decreases monotonically along with the increasing distance from the main stress ($p < .001$). No generalizations can be formulated for words starting with a three-member cluster, as the data features only seven such examples.

Table 4: Tetrasylllabic words in the dictionary (N = 3 683) with initial main stress as a function of initial onset complexity

	Dist 0		Dist 1		Dist 2		Dist 3	
Onset	**No**	**%**	**No**	**%**	**No**	**%**	**No**	**%**
0	144	23	618	40	580	40	24	41
1	414	65	781	50	753	53	34	57
2	77	12	152	10	98	7	1	2
3	1	0	2	0	4	0	–	–
Σ	636	100	1 553	100	1 435	100	59	100

3.2.5. Longer words

The analysis of words composed of five to eight syllables did not reveal uniform results. In pentasyllabics, complex onsets occur in 8% of words with the main stress on the initial syllable, and in 5%, 7% and 9% of words with the stress moved to the left by one, two, and three syllables, respectively (no word starts with a complex onset when the main stress is located on the fifth syllable). In hexasyllabics, words stressed on the first or second syllables do not contain word-initial complex onsets. In the remaining cases, complex onsets occur in 2% of words with the stress on the third syllable, in 8% of words with the stress on the fourth syllable, and in 8% of words with the stress on the fifth syllable. No word-initial complex onsets are attested in hepta- and octosyllabic words.

3.2.6. Token frequencies

This part of the analysis consists in testing the correlations on the corpus data. Token frequencies represent the multiple occurrence of individual words in a large

data resource. The comparison of the results obtained for types and tokens allows us to detect any bias resulting from the disproportionate high frequency of a small group of words. For instance, *draughty* and *grotty* occur 10 and 25 times in the corpus, while the number of occurrences of *question* and *study* amount to 132 806 and 128 326, respectively. Most importantly, the COCA data yields results comparable to the ALDCE data.

First, monosyllabics tend to start with a singleton C (80%). The number of word types decreases gradually from 80% to 0% as the length of word onset increases from C through CCC. Words containing vowel-initial onsets constitute 12% of the monosyllabic words. Second, the results obtained for bisyllabics, trisyllabics and tetrasyllabics confirm that there exists a relationship between the distance from the stressed syllable and onset length. That is, the closer to the main stress, the larger the number of words with heavier onsets. For instance, in bisyllabics with trochaic stress, 22% of words have no onset in the initial syllable, 65% start with C, 12% start with CC, while only 1% contains a CCC-initial cluster. In bisyllabic words with iambic stress, 45%, 52% and 3% of words contain zero, one or two consonants, respectively. Fewer than 1% of lexical items with iambic stress start with a CCC cluster. The distributional patterns for trisyllabics are provided in Table 5. These findings, too, parallel the results obtained for the dictionary data.

Table 5: Percentages of three-syllable words in the corpus (N = 17 770 978) with initial main stress as a function of initial onset complexity

	Dist 0		**Dist 1**		**Dist 2**	
Onset	**No**	**%**	**No**	**%**	**No**	**%**
0	2 245 857	24	3 316 667	43	351 366	60
1	6 031 615	63	3 775 475	49	226 666	38
2	1 144 645	12	575 218	8	13 160	2
3	72 464	1	17 845	0	–	–
Σ	9 494 581	100	7 685 205	100	591 192	100

3.2.7. Complexity of consonant clusters

In the preceding subsections, we have examined the relationship between stress placement and onset complexity in terms of length. We are also interested in the variety of onset types, in particular the consonant clusters that each prosodic position can support. Table 6 lists the number of different C, CC and CCC types found in the dictionary word list. It must be borne in mind that the numbers refer to the presence of a simple onset (1) or a complex onset (2, 3) word-initially.

Table 6: The distribution of onset cluster types across distances

Onset	Dist 0	Dist 1	Dist 2	Dist 3	Dist 4	Dist 5
1	23	22	22	17	13	4
2	46	30	26	17	4	–
3	8	5	3	1	–	–
Σ	77	57	51	35	17	4

Overall, the stressed position (Dist 0) supports the greatest array of cluster types. As the distance from the main stress increases, the number of different C, CC and CCC onset types decreases. As shown in Table 6, the tendency is the strongest in two-member sequences; 46 different types are found in the prosodically strongest position, in contrast to only several types found four syllables away from the primary stress. Distance 6, although reported in section 3.2.5, does not support any consonant clusters, and is therefore excluded from the table.

What is more, the same observations have been made for each individual cluster. Namely, the number of words starting with a specific CC(C) onset decreases from Dist 0 to Dist 5. For example, /pr/ occurs initially in the greatest number of words (Σ = 564), and is supported best in the Dist 0 position (253 words), followed by Dist 1 (227 words), Dist 2 (64 words), Dist 3 (18 words) and finally Dist 4 (2 words). The only exceptions to the pattern are the clusters /wv/ and /skl/, which appear in loans stressed on the second syllable *voyeur, voyeurism* and *sclerosis*.

Accounts of English phonotactics list between 50 and 60 onset or word-initial clusters (Cruttenden 2014; Roach 2006; Trnka 1966; Zydorowicz et al. 2016) depending on the treatment of rare words (e.g. /θj/ in *thurible,* /sv/ in *svelte*) and loans (e.g. /zl/ in *zloty,* /ʃn/ in *schnapps*). As can be observed, stressed onsets (Dist 0) make use of a substantial portion of all the existent clusters (i.e. 54 CC(C)). Similarly, stressed syllables with simple onsets allow 23 out of 24 singleton consonants available in the language (see Roach 2004 and Hillenbrand 2003 for descriptions of the British and American inventories, respectively), with the exception of velar nasals, which are banned from the onset position.

3.2.8. Voicing as a stress predictor

Topintzi (2010) reports on languages in which primary stress is correlated with a voiceless segment in the onset. To test the correlation in English, we inspected the voicing feature of a singleton onset C and the first consonant in CC(C) clusters for the stressed position only (Dist 0). Table 7 summarizes the results, where a voiced segment is labelled with the plus sign (+) and a voiceless segment with the minus sign (–). Table 7 lists the number of segment types with a particular laryngeal feature (Segment) and word types representing them (Word). The data shows that singletons are largely represented by [+voice] consonants, while clusters tend to start with a [–voice] consonant. This result can be attributed to the

phonotactic constraints operating in English (Cruttenden 2014; Roach 2006), whereby the cluster-initial slot tends to be occupied by a voiceless obstruent.

Table 7: Voice contrast in stressed syllables

	C		CC		CCC	
Voice	**Segment**	**Word**	**Segment**	**Word**	**Segment**	**Word**
+	16	6 790	14	861	0	0
-	7	4 372	33	2 676	7	258
Σ	23	11 162	47	3 537	7	258

When we consider the voicing of obstruents, the pattern becomes even clearer. Table 8 presents the number of words starting with voiced and voiceless plosives, fricatives and affricates. To avoid a phonotactically-based bias, calculations are based on singleton onsets only. The results reveal a correlation between the primary stress and voicelessness. That is, onsets with a [–voice] feature outweigh their [+voice] counterparts, conforming the predictions of Topintzi (2010). The tendency holds across all the voiced-voiceless pairs with the exception of /ʧ/ vs. /ʤ/.

Table 8: Voiced and voiceless obstruents in simple onsets

Voiceless C	**Word**	**Voiced C**	**Word**
p	952	b	829
t	670	d	601
k	1 014	g	351
f	677	v	300
s	1 052	z	32
θ	88	ð	11
ʃ	231	ʒ	4
ʧ	200	ʤ	247

4. Discussion

This contribution examines the complexity of onsets in prosodically strong and weak positions, and in relation to their distance towards the stressed syllable. Apart from investigating the binary contrast between stressed and unstressed syllables, which was pursued in previous contributions on the topic, the current study explores the idea of degrees of positional complexity. Our data lead to three major observations.

First, there are a larger number of word-initial onsetless syllables in prosodically recessive positions. For words composed of two, three and four syllables, the number of zero onsets increases monotonically as the distance from the primary stress increases. For instance, in bisyllabics, the proportion between onsetless syllables in the stressed position (Dist 0 = 9%) and in the unstressed

position (Dist 1 = 36%) equals 1:4. This is a static distribution; nevertheless, the pattern is similar to the one attested in Aranda, in which a syllable with an onset is more likely to carry stress than an onsetless one (Topintzi 2010).

Second, the analysis demonstrates an asymmetrical distribution of consonant clusters, with more complex onsets in stressed syllables. Generally, CC and CCC sequences are more likely to occur in the strong syllable. For trisyllabics, 16% of the initial syllables contain CC(C), compared to 10% in the immediately pretonic position and 5% in the position removed from the main stress by two syllables. Quadrisyllabics display the same pattern: stressed syllables contain 12% of complex clusters, and pretonic syllables have 10% (Dist 1), 7% (Dist 2) and 2% (Dist 3) of complex clusters. Also, an inspection of the distribution of different cluster types reveals that pretonic syllables display more varied cluster types than atonic syllables. For instance, stressed syllables host 54 different CC and CCC sequences, in opposition to unstressed syllables.

Finally, the data suggests that, similarly to other languages described in the subject literature, English exhibits a correlation between stress and laryngeal features. Stressed onsets tend to start with a voiceless obstruent. This observation holds not only for singleton segments but also for consonant clusters.

5. Conclusions

Although English has been mentioned among the languages which do not exhibit an interaction of stress with either the onset quality or the presence of an onset, this study reveals that the distribution of onsetless syllables is correlated with the position of stress. First, stressed syllables support not only more complex onsets in terms of length but also a greater array of consonant clusters. Weak syllables, in turn, repel clusters and favour empty onsets or onsets composed of a singleton consonant. Second, the proximity to the stressed syllable also shows to be a predictor of onset complexity; a variety of complex onsets increases as the distance towards the stressed syllable decreases. Overall, the results of the study tip the balance in favour of theories stating that not only the rhyme, but also the onset can play a role in determining the weight of syllabic constituents.

References

Baayen, Harald R., Piepenbrock, Richard and Léon Gulikers. 1993. The CELEX lexical database [CD-ROM]. Philadelphia, PA: Linguistics Data Consortium, University of Pennsylvania.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Clements, George N. and Samuel J. Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.

Cruttenden, Alan. 2014. *Gimson's Pronunciation of English*. 8th edn. London, New York: Routledge.

Davies, Mark. 2011. *Word Frequency Data from the Corpus of Contemporary American English (COCA)*. [Online] Available from: http://www.wordfrequency. [Accessed: 24th January 2011].

Davis, Stuart. 1988. Syllable onsets as a factor in stress rules. *Phonology* 5. 1-19.

Everett, Dan and Keren Everett. 1984. On the relevance of syllable onsets to stress placement. *Linguistic Inquiry* 15. 705-711.

Halle, Morris and Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge, MA: MIT Press.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.

Hillenbrand, James M. 2003. American English: Southern Michigan. *Journal of the International Phonetic Association* 33. 121-126.

Hornby, Albert S. 1974. *Oxford Advanced Learner's Dictionary of Current English*. London: Oxford University Press.

Hyman, Larry. 1985. *A Theory of Phonological Weight*. Dordrecht: Foris.

Kelly, Michael H. 2004. Word onset patterns and lexical stress in English. *Journal of Memory and Language* 50. 231-244.

Orzechowska, Paula, Mołczanow, Janina and Michał Jankowski. 2018. Struktura syllaby a akcent wyrazowy w języku rosyjskim: Badanie korpusowe Studia et Documenta. *Slavica 1–2 (5–6)*, 79-91. DOI: 10.25167/SetDS/2018/1-2/7

Roach, Peter. 2004. British English: Received Pronunciation. *Journal of the International Phonetic Association* 34. 239-245.

Roach, Peter. 2006. *English Phonetics and Phonology: A Practical Course.* Cambridge: Cambridge University Press.

Ryan, Kevin M. 2014. Onsets contribute to syllable weight: Statistical evidence from stressed meter. *Language* 90. 309-341.

Selkirk, Elisabeth O. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Topintzi, Nina. 2010. *Onsets: Suprasegmental and Prosodic Behaviour*. Cambridge: Cambridge University Press.

Trnka, Bohumil. 1966. *A Phonological Analysis of Present-day Standard English*. Alabama: University of Alabama Press.

Zydorowicz, Paulina, Orzechowska, Paula, Jankowski, Michał Dziubalska-Kołaczyk, Katarzyna, Wierzchoń, Piotr and Dawid Pietrala. 2016. *Phonotactics and Morphonotactics of Polish and English: Description, Tools and Applications*. Poznań: Wydawnictwo Naukowe UAM.

AN ANALYSIS OF *CERTAINLY* AND *GENERALLY* IN LATE-MODERN ENGLISH ENGLISH HISTORY TEXTS*

FRANCISCO J. ÁLVAREZ-GIL
Universidad de Las Palmas de Gran Canaria
francisco.alvarez@ulpgc.es

Abstract
This paper analyses the adverbs *certainly* and *generally* as stancetaking markers. These adverbial devices are said to show authorial stance and to communicate the author's commitment or detachment towards the information presented, and so they are classified as epistemic adverbs (Alonso-Almeida 2015). For this study, I have selected a corpus of history texts from the Modern English period (1700-1900), as compiled in *The Corpus of History English Texts* (Crespo and Moskowich 2015), on the basis of which the two evidential adverbs are examined using computer corpus tools, although manual inspection is also employed to assess the meaning of the items in context. The findings suggest that, in this type of scientific articles, the two adverbs are used with differing pragmatic functions, in the case of *certainly* it functions mostly as a booster and, in the specific case of *generally*, its use seems to primarily suggest a hedging purpose (Hyland 2005a).

Keywords: evidentiality, epistemic modality, stance, adverb, corpus linguistics

1. Introduction

This paper follows from ongoing research on the category of adverbs in English in a historical perspective. In this case, I show results concerning evidential adverbs *certainly* and *generally* used as stancetaking markers in the eighteenth and nineteenth century texts in the field of history. Most of these adverbial devices are said to show authorial viewpoint and to communicate either the author's commitment or the author's detachment with regard to the information presented in their context (Alonso-Almeida 2015). This is the reason why adverbials have been chosen as the target linguistic strategies in my analysis, and why they have been considered by most scholars in the field as one of the grammatical categories that most clearly contribute to the expression of interpersonal meanings (cf. Biber and Finnegan 1988, Conrad and Biber 1999). For the research purposes, I have focused on a selection of history texts from the Modern English period (1700-1900), as compiled in *The Corpus of History English Texts* (Crespo and Moskowich 2015), a subcorpus within the *Coruña Corpus of English Scientific*

* Research project: *Etiquetado electrónico de textos científico-técnicos en lengua inglesa entre los siglos XVIII y XX*. Reference: MINECO (FFI2016-75599-P9).

Writing, from where the two evidential adverbs are retrieved using computerised corpus tools, although manual inspection has also been employed.

Even if much research is still to be done in academic texts from a diachronic perspective, there are some previous studies on stance devices carried out within the field of historical pragmatics (cf. Gray et al. 2011, Simon-Vandenbergen and Aijmer 2007, Taavitsainen and Pahta 1997). Following this tradition, I focus on *certainly* and *generally* as examples of a stancetaking feature to see how these forms signal authorial position. My intention is to see whether the analysed adverbs indicate primary epistemic meaning to show degrees of assurance, or there are also some further pragmatic evidential nuances. In my inspection of the instances of the adverbs, I consider the context in which they are embedded. This means especially the syntactic context, since their syntactic distribution may have some important effects on the pragmatics of these and other adverbs, as I have shown elsewhere (Álvarez-Gil 2017). The findings suggest that, in the analysed type of scientific articles, the adverbs are used with differing pragmatic functions; in the case of *certainly* it functions mostly as a booster and, in the specific case of *generally*, its use seems to suggest a hedging purpose in the sense used in Hyland (2005a), as I shall show below.

The structure of the paper is as follows. Section 2 contains the theoretical framework for the analysis of *certainly* and *generally*. This includes the definition of the adverbs as a word-class category along with my own understanding of the concepts of epistemic modality and evidentiality as stancetaking devices. Section 3 is the description of the data analysed as well as the methodology followed in the analysis. Section 4 offers a discussion of the findings. Finally, the conclusions drawn from the present study are presented.

2. Definitions: adverbs, epistemic modality and evidentiality

Adverbs can be considered as a quite heterogeneous word-class category, and this aspect contributes to the difficulty in providing a clear definition of the very category. Van der Auwera (1998), Haspelmath (2001) and Eisenberg (2013), among others, have made explicit reference to this heterogeneity, and they consider this word class

- as being the "most problematic major word class" (Haspelmath 2001: 16543),
- as being an "elusive" (1998: 3) part of speech, and
- as being sometimes "confusing" (2013: 212).

The lack of conceptual unity and the absence of a clear definition of this word-class category and of its scope have led scholars to treat the adverb as an umbrella term for many words which, apparently, are not good candidates for other categories. This can be understood in terms of (a) the diverse morphological structures that adverbs present; (b) the different positions they can occupy within

a clause, and (c) the various syntactic functions they can perform. The resulting problematic situation concerning the misclassification of several words as adverbs has already been highlighted by some scholars (cf. Huddleston and Pullum 2002) who indicated items misclassified syntactically mainly due to formal similarities, including selected cases of adjectives, e.g. *daily*.

Adverbs are often used as stancetaking devices. Stance relates to the expression of the speakers' and writers' "personal feelings, attitudes and value judgements, or assessments" (Biber et al. 1999: 966). Some relevant literature focused on the analysis of stance include Du Bois (2007) and Hyland and Tse (2005). Du Bois (2007: 163) refers to the "stance triangle" when analysing stancetaking in dialogic discourse. The author points out that any act of stancetaking comprises three elements, namely (i) evaluation, (ii) positioning, and (iii) alignment. All this means that whenever we take a stance, we evaluate a certain object, and, at the same time, we position ourselves in an evaluative dimension with respect to that object; and finally, we also associate ourselves with one line of thought. In other words, the stance triangle proposed by Du Bois (2007) emphasises the interactional nature of stancetaking as it necessarily involves specifying (1) where the stancetaker and the subject to whom the stancetaking is targeted stand in relation to the object being evaluated, and (2) where the two subjects stand in relation to each other.

Biber et al. (1999) make a distinction between three main groups of adverbs: (i) circumstance adverbs, i.e. *here*, *now*; (ii) linking adverbs i.e. *nevertheless*, *moreover*, *additionally*; and (iii) stance adverbs, i.e. *really, truly*. Stance adverbs are those that "express the speaker's judgment about the certainty, reliability, and limitations of the proposition; they can also comment on the source of the information" (Biber et al. 1999: 59-60).

In the present study I shall focus on Biber's group three, i.e. the last subcategory. Epistemic stance adverbials (Biber et al. 1999: 59-60) can entail a large number of meanings such as: (a) doubt and certainty, e.g. *perhaps*, *probably*; (b) actuality and reality, e.g. *actually*, *in fact*, *really*; (c) source of knowledge, e.g. *apparently*, *evidently*, *according to*; (d) limitation, e.g. *in most cases*, *typically*, *mainly*; (e) viewpoint or perspective, e.g. *in my opinion*, *from my perspective*; and (f) imprecision, e.g. *kind of*, *roughly*.

On the whole, stance adverbs can be used to indicate the authors' attitude and certainty towards their propositions. The adverbs *certainly* and *generally* are clearly examples of stance adverbs. The form *certainly* can be classified as a perceptual evidential adverb, as it indicates that the evidence the author has for the content s/he expresses has been obtained through the senses. The form *generally*, however, seems to indicate a lower level of authorial commitment to the text content by presenting information in a fuzzier way. These functions represent an evidential and an epistemic value of the two adverbs, respectively.

Evidential devices are "metalinguistic representations of an idea from another source" (Thomas and Hawes 1994: 129) whose function is to indicate the source

of knowledge. Cornillie (2009: 45) takes evidentiality to be a "functional category that refers to the perceptual and/or epistemological basis for making a speech act".

Evidentiality has a strong relation with the concept of stance; however, the latter is classified as an interpersonal feature being defined as the writer's viewpoint on a certain subject matter.

Cornillie (2009: 45) takes evidentiality to be a "functional category that refers to the perceptual and/or epistemological basis for making a speech act". According to Willet (1988), evidentials can be direct, indicating that the speaker/writer has witnessed the action being described, and indirect, indicating that information has been given by someone else or has been obtained through a deductive process. The use of direct evidentials is linked to the acquisition of information through the senses, i.e. visual or non-visual; indirect evidentials, on their part, may belong to the reported category, i.e. first- or second-hand information, or folklore, or to the inferential category.

Evidentiality by many authors is considered to be a subdomain of epistemic modality, in which case it is not only supposed to indicate the writer's source of knowledge, but also some sort of evaluation regarding the truth of the propositional content. This approach can be clearly seen in the works by Chafe (1986), Palmer (1986), and Kranich (2009).

Other scholars hold the position that evidentiality and epistemic modality are two distinct categories (Alonso-Almeida 2015), both semantically and functionally speaking. One case in point is Cornillie (2009: 46-47), who contends that "[e]videntiality refers to the reasoning processes that lead to a proposition and epistemic modality evaluates the likelihood that the proposition is true". This is the position held in this paper, even if secondary pragmatic implications concerning the truth of the proposition may be inferred.

3. The corpus and the method

The *Corpus of History English Texts* (henceforth CHET) is a subcorpus of *The Coruña Corpus of English Scientific Writing* being currently developed and researched at the University of A Coruña. CHET covers a period of two centuries, i.e. 1700-1900. The first text in this subcorpus dates back to 1704, and the last text is from 1895. New patterns of thought, and new methodological procedures based on observation started to be common practices after the seventeenth century (Taavitsainen 2017: 83), and all this is reflected in technical and scientific texts. As regards the genres represented in CHET, there are samples of articles, essays, lectures, textbooks and treatises written by both male and female authors (Moskowich 2017). CHET includes ca. 400,000 words distributed in text samples of 10,000 words. The distribution of words per century is well balanced, 201,794 of the words compiled belong to the eighteenth century, and 202,823 belong to the nineteenth century (Moskowich 2016: 116-117).

My analysis of the texts has been done electronically by means of the *Coruña Corpus Tool* (cf. Moskowich 2016: 118), where the corpus is interrogated by inputting the words *certainly* and *generally* to obtain cases of the adverbs in CHET. In those cases where there is either graphemic or spelling variation, examples are conflated and counted together, as this variation has not proven distinctive from a semantic or a pragmatic standpoint. Statistics are given in the form of raw material, but normalised figures to 10,000 words are offered in the comparison of results according to the variable of sex.

The figures following from the analysis of the corpus are also arranged according to position into left and right periphery. Although the literature abounds with definitions of periphery –all of them referring to position in the clause (cf. Rizzi 1997, Aijmer 2015), I have chosen the orientation in Degand (2014: 154):

> I define the left periphery linearly as the most leftward positional slot of the utterance, outside the dependency structure of the verb. The right periphery lies outside the dependency structure of the verb, at the right of the non-finite verb (if present). Medial position then corresponds to the dependency structure of the clause (predicate-argument structure with adjuncts), which itself has an initial and a final position. The utterance is thus operationalized in clausal terms.

This terminology is similar to the distinction *initial*, *medial* and *final*, as used in Simon-Vandenbergen and Aijmer (2007: 82): "Initial and final refer to the positions respectively preceding and following the other clausal elements. Medial position means that the adverb either precedes (pre-finite) or follows (post-finite) the finite verb directly." Specific positions of a linguistic unit may indicate particular discursive or pragmatic functions. Elements in the periphery, however, are not considered as "part of the syntactic structure of the unit they accompany… They affect discourse conditions, not truth conditions" (Haselow 2015: 160).

4. Results and discussion

The computerised analysis of the text retrieves 46 cases of *certainly* and 76 cases of *generally*. According to the position of the adverbs in the sentence, these appear, as shown in the graph, below (raw numbers):

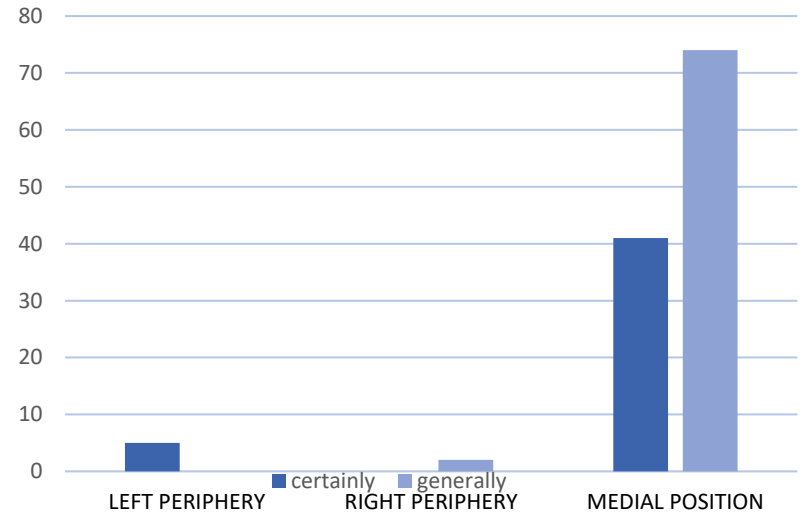


Figure 1. Position of certainly and generally in sentences (raw numbers).

As shown in Figure 1, quantitative analysis reveals that 74 cases of *generally* are given medially in the sentence, while 2 cases are right periphery. As to the form *certainly*, its occurs 41 times medially, and a set of 5 is left periphery. No cases of *certainly* and *generally* have been identified right or left periphery, respectively. As already said, both *certainly* and *generally* appear to emerge more often in medial position in order to modify the meaning of a verb, an adjective or another adverb rather than elsewhere. The adverb *generally* is not registered in CHET, left periphery, and so this form is not used either to be indicative of authorial stance in focus position to modify the complete proposition.

C*ertainly* is also given in the left periphery on few occasions. In the case of *generally*, two examples have been found to occupy a right-periphery position. Considering the variable of sex, as visually shown in Graph 2, *clearly* is used left periphery by men and women indistinctively, but *generally* has been exclusively identified in those texts written by men. The use of *certainly* and *generally* in medial position is also more common in the discourse of men. *Certainly* in medial position is less frequent in texts by women. In the following sections, I review the use of *certainly* and *generally* in the corpus.

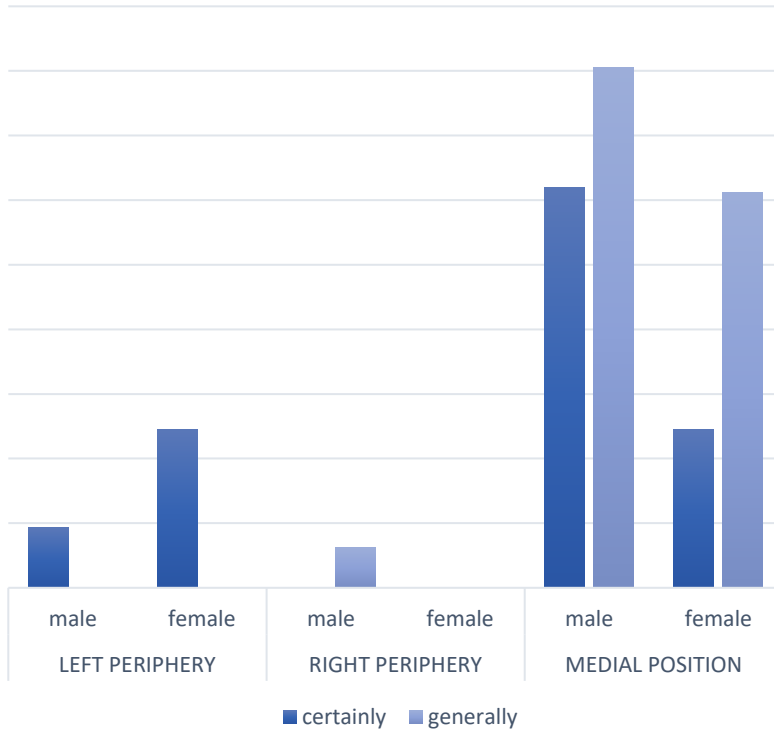


Figure 2. Position of *certainly* and *generally* in sentences according to the variable of sex (normalised figures, N/10,000).

4.1. *Certainly*, left periphery

The adverb *certainly* appears in the texts written by men and women with overuse in the texts by women, as shown in Figure 2, above. Left-peripheral *certainly* holds a dialogic function in that the author's subjective positioning is evinced to indicate authority. This function comes from its metadiscursive potential as a boosting device. At a clausal level, *certainly* designates a consequence logical relationship held between the clause it is in and the previous one.

One instance of *certainly* is given in (1):

(1)

> All strategic considerations were lost in the blaze and glamour of anticipated success; and **certainly** if any one could have accomplished the gigantic enterprise it was Napoleon (hist1895Burrows); female.

This specifies the factual status of the information *it was Napoleon*, which the author truly considers to be the outcome of the logical deduction presented in the

relative construction. The adverb *certainly* comes to reinforce the author's view in the context of the information given previously in the text. This boosting effect is somehow attenuated by the conditioned included in the *if-clause*. This conditional clause includes evidentiary inferential knowledge as entailed in the use of the structure *could* + perfective 'could have accomplished' indicating the author's subjective elaboration of information.

The boosting effect of *certainly* seems stronger in (2), and, from a discourse perspective, *certainly* suggests a consequence relationship between the two neighbouring sentences:

(2)

> He was the first clergyman of the establishment appointed to the parish, and is also stated to have been the first of the body that officiated in the parish church, as tradition asserts, that, before his arrival, the Presbyterians used it as their place of worship; and, **certainly**, the building, in its primitive state, was well suited to their taste, as it had neither spire, belfry, stained-glass windows, nor any of those other ornaments which the Puritans regarded as sinful and ungodly (hist1862Bennett); male.

The use of *certainly* in this instance follows from the author's subjective positioning. The information provided earlier in the text cannot be safely attributed to any identifiable source; yet the author relies on an intersubjective source of information, as identified in the use of "as tradition asserts". In this context, responsibility for the claims made in this passage is collective, even if the use of *certainly* may indicate the author's conclusion. If the instances in (1) and (2) are compared, the former shows a nervier attempt to show information, even if protection of her face only relies on the logical estimate put forward by the use of the conditional. The latter appears to depart from giving a bold expression of self-accountability in his elaboration of information, and the idea of keeping his public face intact seems to lie behind the use of *certainly* initially.

This same reason justifies the use of *certainly* in front position to stress the validity of the complete proposition in (3). The evidential nature of this adverb is explained in the information given back in the text, which presents both evidential and evaluative devices, namely *knowing from his own experience* and *he judged rightly*, respectively.

(3)

> Beſides employing the poor, he was much ſet on inſtructing them in religion, well knowing from his own experience, that piety is the foundation on which all other virtues muſt be built. And in that he judged rightly. For **certainly** thoſe are moſt likely to ſubmit to the evils of life with patience, and to fulfil the duties of it amidſt temptations and ſnares,

who have a firm faith in, and a good hope towards God (hist1780Cornish); male.

The strengthening validity of the evidential device *certainly* contrasts with the use of *generally*, which appears to be used with a clear mitigating effect, regardless its position in the sentence.

4.2. *Generally*, right periphery

The adverb *generally* given finally in the sentence only occurs in those texts written by men. As explained earlier in section 3, the use of adverbials in this position has an effect on the discourse conditions of the text, rather than on the truth conditions. In all the cases of *generally* given in the right periphery seems to be indicative of an intersubjective authorial position with respect to what they describe. The use of the adverb *generally* pursues a mitigating effect, as it avoids offering information accurately. This hedging function of generally could be classified as an example of vague language (Cutting 2007: 6): "language that is inherently and intentionally imprecise". Imprecision in the instances detected in our corpus has a source in the authors' desire not to be assertive (Carter and McCarthy 2006: 202). Some examples taken from the corpus are included in (4) and (5):

(4)

The ſtrength and vivacity of expreſſion, the beauty and variety of thought are almoſt inimitable. This excellent author paſſes over the tranſactions of the two firſt propraetors under *Veſpaſian* ſlightly and **generally**, but is more full and particular in his relation of what paſſed during the command of *Agricola* (hist1732Horsley); male.

(5)

In connexion with the history of Tara at this period, two very important events are stated by the modern historians **generally** (hist1839Petrie); male.

In these examples, the use of *generally* appears to suggest both lack of details of the information passed on to them from earlier and contemporary historians. The adverb has here a hedging function to avoid full commitment with the truth of the events mentioned in (4) and (5).

4.3. *Certainly* and *generally* in medial position

The major presence of the adverbs under analysis is found medially in the sentence, as has been pointed out at the beginning of the section. The form *generally* is the most frequent device in CHET used as a premodifier preceding

the verb, with a higher frequency in those texts written by women, as revealed in Table 1.

Table 1. *Certainly* and *generally* in mid-position (normalised figures, N/ 10,000).

	M	F	M	F	M	F	M	F	M	F	M	F
	ADV + V		V + ADV		BE + ADV		ADV + BE		ADV + ADJ		ADV + ADV	
certainly	0.50	0.37	0.06	0.00	0.19	0.00	0.03	0.00	0.09	0.00	0.25	0.24
generally	0.99	1.47	0.15	0.12	0.09	0.00	0.09	0.12	0.40	0.00	0.06	0.37

In this chart, figures concerning other modifying uses of these adverbs are offered. These figures also indicate that *certainly* as a postmodifier does not occur after a lexical verb (V) or any form of BE in texts by women, while it shows in texts by men. Women do not use these adverbs as a premodifier with the obliques of BE or with adjectives. In general, this table and the graphs already shown suggest that differences do not really apply to the distribution of these two adverbs. It shows, however, that there is a preference for *certainly* in the texts written by men.

The following instance of *certainly* in preverbal position in (6), the adverb is used to show the author's conviction towards the information he owns. This use of *certainly* is contextually supported by the direct visual evidential *I have ſeen* given later in the text; in this way, the author shows both his involvement in the elaboration of information, while also arguing for his own responsibility in making this claim.

(6)

> As to the preciſe time of *Barſcube's* Deſcent, I cannot determine; but this much I **certainly know**, that they were a younger Son of that Noble Family: For I have ſeen a Charter granted by *Matthew* Earl of *Lenox, dilecto conſanguineo ſuo Thomæ Stewart de Terris de North-Bar, Craigtoun, Barſcube; Raſhielee, apud Crocſtoun 5. Julii Anno* 1497 (hist1710Crawfurd); male.

The same sense of conviction is clearly noticed in the use of certainly in (7), and so this adverb reinforces the conclusion reached at after the deductive process given earlier in the text. This deductive process is specified in the cotext by the the evidential devices might and could followed by the perfective: might have been and could have remembered. The reasoning process is also supported by (a) the use of lexical items showing interpretation and degress of subjectivity, namely speculate and the result of and (b) by the logical procedure established in the conditional sentence:

(7)

> It is not of much use to speculate as to what might have been the result of a Peace in 1797. If the French could have remembered the ancient maxim that „half is better than the whole," they would **certainly** have left off in their career of conquest at a much better point for the extension of their empire and of their political principles than they secured at the Treaty of Vienna (hist1895Burrows); female.

The use of *certainly* as a postmodifier in mid position after the verb shows the stance of the author and represents evaluative rather than evidential meaning, and this is evinced in (8), below. The adverb in this example comes to be revealing of the author's appraisal of the way the Earl of Sussex should understand the event described.

(8)

> Queen Elizabeth, in Scotland, having by his letters from Edinburgh, of the 19th of Auguſt, given ſome occaſion to the Engliſh court to doubt, whether the Baptiſm might not be ſolemnized on the Sunday following, though he had advertiſed the King on what day the Earl of Suſſex began his journey; the Lord Burghley, Lord High Treaſurer, wrote to that Earl a letter, the original of which is ſtill extant in the Cotton library, *Calig. D.* II. dated from the court at Greenwich the 28th of Auguſt 1594, acquainting him, that if he ſhould underſtand **certainly**, that the ceremony of Baptiſm was over, her Majeſty's pleaſure was, that he ſhould not go into Scotland, nor ſuffer her preſent to be carried thither; but that he ſhould, as of his own diſcretion, without ſeeming to have direction from the court, advertiſe Mr. Bowes, that he, the Earl, meant not to come into Scotland, conſidering, that his meſſage was only to be preſent at the Baptiſm (hist1760Birch); female.

The adverb *certainly* occurring with any form of BE can show after and before this verb, as exhibited in (9) and (10):

(9)

> It must, however, on the other hand, be carefully noticed, that though the Bill of Rights might not propose itself as any alteration, it was **certainly** a complete renovation of the free constitution of England; the abject state to which the laws, the constitution, and the people themselves, had fallen, must never be forgotten (hist1840Smyth); male.

(10)

> But though the bait was visible to people of any sense, to the multitude it **certainly** was in no small degree alluring (hist1800Stock); male.

The position of the adverb in this respect indicates different perspectives, but both instances reinforce authorial position in relation to the statements made. The use of *certainly* in (10) reports on a more solid position and evaluation of the writer towards the information presented and so he shows his entire opposition. The use of this adverb in (9) suggests a less involved authorial stance than the case in (10).

A strengthening effect is also seen in the use of *certainly* preceding adjectives, as in (11) and (12), in which the author makes clearly manifest his position towards the information presented in (11) and an evaluation of the degree of truthfulness of the account in (12). The adverb in both cases can be subsequently categorised as a booster in the sense in Hyland (2000: 179). For Hyland, boosters "are words such as *clearly*, *obviously* and *demonstrate*, which allow writers to close down alternatives, head off conflicting views and express their certainty in what they say" (Hyland 2005b: 52). The use of *certainly* in (11) and (12) below fits within this definition.

(11)

There is **certainly** little in these transactions to countenance any experiments or reasonings of this nature (hist1840Smyth); male.

(12)

He is often censured and sometimes corrected, but the fact seems to be, that without his original, and **certainly** honest account, we should know little about the events and affairs he professes to explain (hist1840Smyth); male.

Finally, mid-position *certainly* preceding an adverb has a clear strengthening effect, and so it might be used to emphasize the meaning of the adverb, as in (13) in which the adverb is used as an intensifier preceding the adverb of negation *not*. The same is true in the case of *certainly* in (14) used as a premodifier to increase the illocutionary potential of the adverb *not*, and so the evidential value of the adverb *certainly* is noticeable. So much so that *obviously* and *evidently* can be synonyms for *certainly* in these contexts:

(13)

But Livy, though one of the moſt beautifull, is **certainly** not the moſt exact hiſtorian (hist1750Chapman); male.

(14)

No incident of the whole life of Laud has exposed him to severer comment than this, and **certainly** not without reason (hist1833Aikin); female.

In the case of *generally*, the use of this adverb in combination with evidential forms, either of a cognitive nature, such as *know* and *understand* in (15), (16) and (17), or of a communicative nature, such as *say* in (18), is interesting. It seems that

the illocutionary force of these evidential verbs is somehow softened by the presence of *generally*, which suggests some vagueness in authorship. In other words, it reports on an unspecified conceptualiser, and consequently attribution is opaque. This means that responsibility for the claims made is shared; hence, *generally* is here representative of an intersubjective device.

(15)

The odious character he had, upon account of his cruelty and lewdneſs, is **generally known** (hist1732Horsley); male.

(16)

It is not, perhaps, **generally known** that our annals, simple as they may be, reach back nearly two centuries, and are connected in their origin with one of the most remarkable religious orders the world ever saw, a brief notice of which seems to be necessary (hist1884Breese); male.

(17)

In short, neutrality was impossible, both on social and political grounds; and this is now very **generally understood** (hist1895Burrows); female.

(18)

Whilst confined in Carisbrooke castle, it is **generally said** that the king wrote the *Eikon Basilike*; or, Portraiture of his own Sufferings (hist1814Britton); male.

The presence of *generally* as a modifier of adjectives and other adverbs implies some degree of vagueness in discourse in examples (19), (20) and (21), namely *generally monotonous* and *generally great* (ADV+ADJ), and *generally utterly* (ADV+ADV). The intentional use of *generally* to convey vague language in these texts might pursue a mystifying effect, and this emerges from a desire to reduce authorial responsibility in making the claim. Opposite to the use of *certainly* described above, *generally* could be also considered as a negative politeness strategy (Brown and Levinson 1987) to avoid full imposition on readers, while, at the same time, this strategy reduces the chances of a face-threatening act.

(19)

This northern front however is **generally** monotonous in effect (hist1814Britton); male.

(20)

Another of their Amuſements is in the Week before *Lent*; which the *Ruffians* call the *Butter-Week*. They go about three Miles out of Town, where the Hills are of great Eminence; and both Men, Women, and Children, divert themſelves by ſliding down in Sledges. Some will do it in

a moſt dexterous Manner. But there are **generally** great Numbers, who break their Legs, and Arms. This Diverſion is ſo highly honoured, as to have Her Majeſty for a Spectator (hist1739Justice); female.

(21)

And it was nothing but the want of union among themſelves that reſtrained the ſavages from breaking up all our ſettlements, which were **generally** utterly defenceleſs (hist1769Adams); male.

In brief, there is a clear distinction between the use of *certainly* and *generally* in mid-position. Their use in the scientific texts analysed reveals that both adverbs are used to clearly show authorial stance. While *certainly* has always a strengthening function, *generally* fulfils a mitigating role. The former is used subjectively, while the latter is used intersubjectively. The form *certainly* is often combined with evidential devices to support the authors' perspectives, and so they seek for some kind of partnership and support in making a claim. The item *generally* indicates reliability on ther external sources, but without really committing to them.

From a gendered perspective, the analysis of *certainly* and *generally* right periphery, left periphery and medial in the clause shows no major difference between men and women in the sense that all of them pursue a desire to use mitigating resources to hedge their claims. Mondorf (2002: 176) reports, however, that women are often said to deploy more attenuating devices than men. Leaper and Robnett (2011) conclude the opposite, and so they hold that "tentative speech should be viewed as both women's *and* men's language" (2011: 139). Evidence collected here, albeit a small amount, comes to confirm Leaper and Robnett's view, although I have shown that women tend to include contextual hedges in the vicinity of assertive claims with *certainly*. Still, this is an earlier observation in the light of the data provided in CHET, and should be checked against data obtained from a much larger and multigenre corpus.

5. Conclusion

In this paper, my overall aim has been to evaluate the adverbs *certainly* and *generally* as used in context in early scientific writings. For this, I have used samples excerpted from the discourse of history texts in late Modern English. It seems that the syntatic position of these adverbs rarely affects their meaning. It is true, though, that the adverbs *certainly* and *generally* do not occur in all positions in the text analysed. The term *certainly* does not happen in the right periphery position, and the item *generally* is not registered in the left periphery in CHET.

From a pragmatic standpoint, results support my initial expectations, and they indicate that *certainly* follows from evidential implications contextually triggering the use of this adverb to reinforce the author's position concerning the information

presented regardless of their syntactic distribution. On the opposite side, there is the use of *generally*, which seems to suggest the author's intention to reduce responsibility by invoking imprecision in language. Likewise, it has been shown that the use of *generally* with evidential lexical verbs serves to downtone the illocutionary force of the proposition either by indicating shared responsibility or by claiming unidentified attribution.

The adverbs *certainly* and *generally* seem to show a dialogic function in the texts analysed, as they help to accommodate the author's viewpoint in their academic community following the tradition of linguistic politeness to avoid imposition. The pursuit for self-protection of their public image is also in the balance, and these adverbs can help in this matter because they are deployed intersubjectively. Otherwise, the form *certainly* would imply a strong deontic force, unless contextually attenuated by the presence of other justificatory claims.

Concerning the variable of sex, this has been a secondary issue in this paper, and much research would be needed in this respect with a larger body of data, as already said in this paper. I have shown that, while all my authors use *certainly* and *generally* with the same discourse and pragmatic functions from a gender perspective, the texts written by women tend to present patterns of the adverbs with a mitigating function supported by contextual attenuating devices. They also include fewer examples of the anaysed adverbs with strengthening the function of claryfying their stance. This may show the imbalance existing between men and women scientists, as women had to show themselves in a different hierarchical position in the process of elaboration of meaning. The position implies a higher degree of modesty in an attempt to reach wider acceptance within their male-dominated scientific community. Yet, a full gendered approach on a larger database is still in order as future research to detect possible variation concerning the patterns of tentative language as deployed by men and women along with the definition of the frequency of this type of gendered language patterns.

References

Aijmer, Karin. 2015. Analysing discourse markers in spoken corpora: *actually* as a case study. In Paul Baker and Tony McEnery (eds.), *Corpora and Discourse Studies Integrating Discourse and Corpora*, 88-109. London: Palgrave MacMillan.

Alonso-Almeida, Francisco. 2015. Introduction to stance language. *Research in Corpus Linguistics,* 3. 1-5. Available at: http://www.aelinco.es/ojs/index.php/ricl

Álvarez-Gil, Francisco J. 2017. *Apparently*, *fairly* and *possibly* in *The Corpus of Modern English History Texts* (1700-1900). In Francisco Alonso-Almeida (ed.), *Stancetaking in Late Modern English Scientific Writing. Evidence from the Coruña Corpus.* Colección Scientia [Applied Linguistics]. Valencia: Servicio de Publicaciones de la Universidad Politécnica de Valencia.

Biber, Douglas and Edward Finegan. 1988. Adverbial stance types in English. *Discourse Processes* 11 (1). 1-34.

Carter, Ronald and Michael McCarthy. 2006. *The Cambridge Grammar of English*. Cambridge: Cambridge University Press.

Chafe, Wallace. 1986. Evidentiality in English conversation and academic writing. In Wallace Chafe and Johanna Nichols (eds.), *Evidentiality and the Linguistic Coding of Epistemology*, 261-272. Norwood, NJ: Ablex.

Conrad, Susan and Douglas Biber. 1999. Adverbial stance marking in speech and writing. In Susan Hunston and Geoff Thompson (eds.), *Evaluation in text: Authorial stance and the construction of discourse*, 56-73. Oxford: Oxford University Press.

Cornillie, Bert. 2009. Evidentiality and epistemic modality. On the close relationship between two different categories. *Functions of Language* 16 (1). 44-62.

Crespo, Begoña and Isabel Moskowich. 2015. A Corpus of History Texts (CHET) as part of the Coruña Corpus Project. In *Proceedings of the international scientific conference*. St. Petersburgh State University, 14-23.

Cutting, Joan. 2007. *Vague language explored*. New York: Palgrave.

Degand, Liesbeth. 2014. 'So very fast very fast then' Discourse markers at left and right periphery in spoken French. In Kate Beeching and Ulrich Detges (eds.), *Discourse functions at the left and right periphery: Crosslinguistic investigations of language use and language change*, 151-178. Brill: Leiden.

Du Bois, John W. 2007. The Stance Triangle. In Robert Englebretson (ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction,* 139-182. Amsterdam: John Benjamins.

Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik. Der Satz*. 4th edn. Stuttgart: Metzler.

Gray, Bethany, Douglas, Biber and Turo Hiltunen. 2011. The expression of stance (1665-1712) in early publications of the Philosophical Transactions and other contemporary medical prose: Innovations in a pioneering discourse. In Irma Taavitsainen and Päivi Pahta (eds.), *Medical writing in Early Modern English*, 221-247. Cambridge: Cambridge University Press.

Haselow, Alexander. 2015. Left vs. right periphery in grammaticalization The case of anyway. In Andrew D.M. Smith, Graeme, Trousdale, and Richard Waltereit (eds.), *New Directions in Grammaticalization Research*, 158-186. Amsterdam/ New York: John Benjamins.

Haspelmath, Martin. 2001. Word classes/parts of speech. In Neil Smelser and Paul Baltes (eds.), *Encyclopedia of the Social and Behavioral Sciences*, 16538-16545. Oxford: Pergamon.

Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press.

Hyland, Ken and Polly Tse. 2004. Metadiscourse in academic writing: a reappraisal. *Applied Linguistics*, 25 (2). 156-177. doi.org/10.1093/applin/25.2.156

Hyland, Ken. 2000. Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language Awareness,* 9 (4). 179-197. doi.org/10.1080/09658410008667145

Hyland, Ken. 2005a. Stance and engagement: A model of interaction in academic discourse. *Discourse Studies,* 7 (2). 173-192. doi: 10.1177/1461445605050365

Hyland, Ken. 2005b. *Metadiscourse*. London/New York: Continuum.

Kranich, Svenja. 2009. Epistemic modality in English popular scientific texts and their German translations. *Trans-kom,* 2 (1). 26-41. https://d-nb.info/999839527/34

Leaper, Campbell and Rachael Robnett. 2011. Women are more likely than men to use tentative language, aren't they? a meta-analysis testing for gender differences and moderators. *Psychology of Women Quarterly*, 35 (1). 129-142. doi.org/10.1177%2F0361684310392728

Mondorf, Britta. 2002. Gender Differences in English Syntax. *Journal of English Linguistics*, 30 (2). 158–180. https://doi.org/10.1177/007242030002005.

Moskowich, Isabel. 2016. Lexical richness in modern women writers: Evidence from the Corpus of History English Texts. *Revista canaria de estudios ingleses,* 72. 111-128.

Moskowich, Isabel. 2017. Genre and change in the Corpus of History English Texts. *Nordic Journal of English Studies,* 16 (3). 84-106.

Palmer, Frank R. 1986. *Mood and modality* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Liliane Haegeman (ed.). *Elements of Grammar*, 281-337. Dordrecht: Kluwer.

Simon-Vandenbergen, Anne-Marie and Karin Aijmer. 2007. *The semantic field of modal certainty. A corpus-based study of English adverbs*. Berlin/New York: Mouton de Gruyter.

Taavitsainen, Irma and Päivi Pahta. 1997. Corpus of early English medical writing 1375- 1550. *ICAME Journal,* 21. 71-78.

Taavitsainen, Irma. 2011. Medical case reports and scientific thought-styles. *Revista de lenguas para fines específicos,* 17. 75-98.

Thomas, Sarah and Thomas P. Hawes. 1994. Reporting verbs in medical journal articles. *English for Specific Purposes*, 2. 134-155.

Van der Auwera, Johan. 1998. Introduction. In Johan van der Auwera (Ed.) Adverbial Constructions in the Languages of Europe. Berlin: De Gruyter, 1-24.

Willet, Thomas. 1988. A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language,* 12. 51-97.

SPECIALIZED KNOWLEDGE REPRESENTATION: FROM TERMS TO FRAMES*

PAMELA FABER
University of Granada, Spain
pfaber@ugr.es

MELANIA CABEZAS-GARCÍA
University of Granada, Spain
melaniacabezas@ugr.es

Abstract
Understanding specialized discourse requires the identification and activation of knowledge structures underlying the text. The expansion and enhancement of knowledge is thus an important part of the specialized translation process (Faber 2015). This paper explores how the analysis of terminological meaning can be addressed from the perspective of Frame-Based Terminology (FBT) (Faber 2012, 2015), a cognitive approach to domain-specific language, which directly links specialized knowledge representation to cognitive linguistics and cognitive semantics. In this study, context expansion was explored in a three-stage procedure: from single terms to multi-word terms, from multi-word terms to phrases, and from phrases to frames. Our results showed that this approach provides valuable insights into the identification of the knowledge structures underlying specialized texts.

Keywords: context expansion, frame, multi-word term, phrase, specialized discourse

1. Introduction

An important issue in translation is how to achieve sameness of meaning across languages and at all levels of the text. In the case of the translation of scientific and technical texts, a considerable percentage of translation quality depends on finding optimal correspondences for the specialized language units or terms used to convey the text message. These units, which may be single or multi-word terms, designate objects, events, processes, and attributes in the specialized field (Faber 2012).

Terms, semantic clusters of terms, and their configurations activate segments of the conceptual structure of a knowledge domain (Sager et al. 1980), which is present in the source and target language-cultures. If both language-cultures have

* This research was carried out as part of FFI2017-89127-P, *Translation-oriented Terminology Tools for Environmental Texts (TOTEM)*, research project funded by the Spanish Ministry of Economy and Competitiveness. Funding was also provided by an FPU grant given by the Spanish Ministry of Education to the second author.

terms for the entities designated, the assumption is that the text can be translated with a reasonable degree of accuracy. The translator must first be aware of what is happening in the text and the message that it conveys. Then he/she identifies term correspondences and finds the most accurate way to link them so as to highlight the semantic relations between concepts that are explicit in the text.

Understanding specialized discourse thus depends on the text receiver's capacity to grasp and then activate the knowledge structures underlying the text. When the text receiver is not an expert in the field (as often occurs in a specialized translation scenario), he/she must be able to rapidly acquire the necessary domain-specific knowledge (Faber 2012).

In the translation process, the specialized knowledge units in a text as well as their relations must be analyzed at various levels. Although the meaning of certain concepts and relations are evident in the surface structure of the text, this is merely the tip of the iceberg. There is a whole world of meaning lurking beneath the surface, which translators must be able to perceive. Relevant data from the source language text must be generalized or abstracted with a view to integrating new information into semantic memory. Understanding thus depends on the translators' ability to successfully construct a mental representation of a segment or segments of the specialized knowledge field. The expansion and enhancement of knowledge is thus an important part of the specialized translation process (Faber 2015).

This paper explores how the analysis of terminological meaning can be addressed from the perspective of Frame-Based Terminology (FBT) (Faber 2012, 2015), a cognitive approach to domain-specific language, which directly links specialized knowledge representation to cognitive linguistics and cognitive semantics. In FBT, knowledge acquisition involves a progressive expansion of meaning, which begins at the term-level, progresses to the phrase level, and finally results in the codification of an entire knowledge frame.

2. Theoretical background

To understand how knowledge is configured, and expanded, it is necessary to start with the brain. Neurological studies provide insights into how experts retrieve and activate stored knowledge (Quillian 1969; Anderson 1983; Gallese and Lakoff 2005; Patterson et al. 2007; Meteyard et al. 2012; Kiefer and Pulvermüller 2012). For this reason, Faber et al. (2014) conducted a pilot fMRI study in which brain activation images of expert geologists were compared to those of novices as they performed a series of different tasks, such as linking geological tools to their function and tools to images. The results showed that expert knowledge involves a supramodal conceptual representation, which transcends sensory input modalities such as vision or hearing. Conceptual representations thus have

multiple levels of input (Binder and Desai 2011), which do not only come from the senses.

At the top level, much research agrees that there is a non-modality-specific schematic representation, which is progressively fleshed out by sensory-motor-affective input when and as needed (Patterson et al. 2007). Faber et al. (2014) highlighted the key role played by contextualization and situation in specialized knowledge processing since the brain regions activated by experts (though not novices) were those strongly implicated in mental imagery, episodic memory, and context representation. Although more studies are necessary, Faber et al. (2014) further validated the need to explore how contextual information can be activated and thus facilitate frame creation in the non-expert.

Accordingly, FBT applies the notion of 'frame' (Minsky 1975; Fillmore 1985, 2006) defined as a schema or knowledge structure, which relates elements and entities associated with a particular scene, situation that is part of human experience. A frame is thus as an organized package of knowledge that humans retrieve from long-term memory to make sense of the world (Faber 2012). Given that concepts cannot exist in a vacuum, they are more meaningful when they are related to each other and integrated into progressively more complex knowledge configurations. Framing experience involves applying stored knowledge derived from similar contexts and situations with a view to understanding complex events and how to deal with them.

In Terminology, the usefulness of embedding concepts in situations has also been highlighted as a way of enriching conceptual representations (Dubuc and Lauriston 1997; Faber 2012; Temmerman 2013). Although context is often regarded as the segment/s which precede or follow a word or phrase (Lyons 1995), it can also be a related situation, events, or information that help users to understand something, and which reflect a specific knowledge profile (Kecskes 2014; Faber and León-Araúz 2016). The specification of contexts should thus take place at multiple levels that range from concept to frame.

3. MWTs and context expansion

As is well known, MWTs are terms composed of more than one word. In English, they can be of varying length: (i) two constituents (*transboundary pollution*); (ii) three constituents (*surface water pollution*); (iii) four constituents (*wood-burning-stove pollution*); and even (iv) five constituents (*volatile organic compounds pollution source*). Although in theory, MWTs can go on forever, it is extremely rare to find one longer than four or, at most, five words because of the cognitive demands made on text receivers.

In English, these complex terms resemble a type of expert shorthand, where there is no need for further explanation because of the level of knowledge presumably shared by the text sender and receivers. Users are thus obliged to unpack the meaning of MWTs and correctly access the relationship between the

constituents. To do this, they must mentally activate a specialized event or frame in which the relations between participants are specified. Although this is relatively easy for an expert in the field, it can be somewhat more difficult for a non-expert.

Consequently, the process of understanding terminological meaning begins with the concept designated by the term itself and is conceived as a progressive expansion of contexts. First, there is the term and its microcontext (Cabezas-García and Faber in press), which can be further expanded to a set of related multi-word terms (MWTs). As shall be seen, these MWTs can subsequently be unpacked by inserting implicit information and then by explicitly relating them to each other as well as to other specialized knowledge units. As we shall see, this gives rise to the specification of larger knowledge structures or frames.

3.1. From single terms to MWTs

Context expansion initially takes place when a single term undergoes further specification and becomes a multi-word term (MWT). In specialized language, most MWTs take the form of endocentric noun compounds (Nakov 2013), e.g. *climate change*.

Endocentric MWTs are informative because they point to relations between and within semantic categories. Generally speaking, an endocentric MWT is a specialization of the meaning of its head. This means that term structure can often be used as a way to automatically extract information regarding conceptual hierarchies as well as hyponymy subtypes (Sager et al. 1980). For example, *vessel-source marine pollution,* is a type of *marine pollution,* which is a type of *pollution.* For further semantic characterization, we can also say that pollution *affects* the sea and *is caused by* vessels.

In morphologically-poor languages, such as English, endocentric MWTs can take the form of sequences or stacks of nouns of varying length. In English, lengthy pre-modification in the form of a series of nouns, also modified by adjectives or even entire phrases, is a frequent method that is used to condense and concentrate domain-specific knowledge (Sager et al. 1980; Štekauer et al. 2012; Fernández-Domínguez 2016).

Concept specialization involves a slot-filling mechanism where the modifier is inserted into a slot in the head-noun schema, also known as its micro-context (Cabezas-García and Faber in press). In an MWT, the modifier is directly related to the base meaning of the head noun as (under)specified in the definition and is interpreted accordingly. In the second stage, world knowledge is used to expand the context of the headword and its interpretation.

For example in the case of *pollution,* this expansion of context starts with its definition:

(1)

> ***pollution*** presence in the **environment** [Slot 1] of a **substance** [Slot 2], whose **nature** [Slot 3], **source** [Slot 4], **location** [Slot 5] or **quantity** [Slot 6] produces **undesirable effects** [Slot 7] for the environment or the health of living organisms.

The general concept of *pollution* is thus defined in terms of seven meaning slots: (i) environment; (ii) substance; (iii) nature; (iv) source; (v) location; (vi) quantity; and (vii) undesirable effects. All of these slots are underspecified and thus susceptible to be filled by hyponyms of the terms in bold. When one or more of these slots are made more specific, this generates MWTs that are hyponyms of *pollution.* Table 1 shows examples of sets of MWTs corresponding to each slot.

Table 1. MWT hyponyms of pollution

Definition slots: *pollution*	**Multi-Word Terms**
ENVIRONMENT	**air** pollution, **water** pollution, **soil** pollution, **marine** pollution, **ocean** pollution
SUBSTANCE	**oil** pollution, **particle** pollution, **solid waste** pollution, **nutrients** pollution
NATURE (of substance)	**volatile organic compounds** pollution
SOURCE (of substance)	**point-source** pollution, **non-point-source** pollution, **wood-burning-stove** pollution, **industrial** pollution, **traffic-related** air pollution
LOCATION (of substance)	**transboundary** pollution, **transfrontier** pollution,
QUANTITY (of substance)	**intensive** air pollution
UNDESIRABLE EFFECTS	**oxygen depletion** pollution, **thermal** pollution

As can be observed in Table 1, this underspecified meaning of *pollution* is a rich source of possibilities since it predicts the subclasses of MWTs that can designate more specific types of pollution. This allows translators to grasp the different dimensions of pollution or perspectives from which the pollution process can be envisaged.

Knowledge of the types of entity that can fill those slots facilitates understanding of MWTs. This is important because in such cases, syntax cannot be used to clarify meaning. This is evident in compounds such as *water pollution* and *oil pollution.* Despite the fact that *water pollution* and *oil pollution* possess the same syntactic structure (N+N) and even combine the general semantic categories of LIQUID and PROCESS, they obviously differ in the semantic relation between modifier and head, as reflected in their definition slots. This means that

water is the affected entity or patient of pollution, whereas *oil* is the polluting agent (Cabezas-García and León-Araúz 2018).

Even though *water* and *oil* belong to the same semantic category of LIQUID, the accurate interpretation of *water pollution* and *oil pollution* depends, among other things, on the conceptual distinction between INGESTIBLE LIQUID and NON-INGESTIBLE LIQUID as well as the functions of both. *Water,* which is ingestible and necessary for life, is highly sensitive to pollution. In contrast, *oil,* which is non-ingestible and used as a fuel, can have a negative impact on water since oil is a polluting agent that destroys marine life.

This is a basic example of the general knowledge that users must be able to access and activate for an accurate interpretation of both terms. Having this information available at some level signifies that at least a partial representation of semantic structure must be encoded, and enriched by pragmatic information. Syntax and surface form is not sufficient (Štekauer et al. 2012; Buendía Castro and Faber 2016).

For example, in scientific and technical translation from English into another language, the translator does not generally possess the same level of expert knowledge as the source-language text receivers. When the translation is from English into morphologically-rich languages such as Spanish or French, where noun-stacking is not an option, it is necessary to make the relations between MWT components explicit, usually in the form of adjective or prepositional postmodification (Maniez 2009; Daille 2017). In the case of Spanish, the translation of *water pollution* would be *contaminación* ***del*** *agua* whereas *oil pollution* would be translated as *contaminación* ***por*** *hidrocarburos*. The prepositions *de* [of] and *por* [by] are used to encode the conceptual relations implicit in the English MWT.

3.2. Multi-word term level to phrase level

Multi-word terms (MWTs) are also characterized by concealed propositions that can be inferred in the term-formation processes (Levi 1978). This means that MWTs can also be further expanded, especially since many of these terms are the result of predicate deletion (*transboundary pollution* instead of 'pollution crosses boundaries') or predicate nominalization (*chemical water pollution* instead of 'chemicals pollute water'). Both of these term-formation processes have predicate-argument structure.

As is well known, 'predicate argument structure' refers to the lexical representation of argument-taking lexical items (Levin 2013). These are typically verbs and their nominalizations. The specification of argument structure involves identifying the number of arguments that a lexical item can take, their syntactic expression, and their semantic relation to the predicate.

Although syntactic expression is language-specific, semantic relations are not. Semantic relations can be understood as semantic roles such as AGENT, PATIENT,

INSTRUMENT, EXPERIENCER, LOCATION, etc. Although most linguists tend to believe that semantic roles are a good idea, at least in some form, there is considerable disagreement as to their number, nature, and function. Currently, there are as many inventories of semantic roles as there are theories that use them (Van Valin and LaPolla 1997; Gildea and Jurafsky 2002; Fillmore et al. 2003; Palmer et al. 2005).

If we take a look at the argument structure of *pollute,* it would have the same number and semantic type of arguments as its correspondences in different languages (i.e. *polluer, verschmutzen, contaminar, inquinare, polua,* etc.). In all language-cultures, *pollute* is characterized by a polluting agent as well as a polluted (or affected) entity. The propositional representation of *pollute* is thus a type of *tertium comparationis* that can be used as the basis for semantic equivalence (Buendía Castro and Faber 2016). In fact, this type of representation and information is used, at least in some form, in various machine translation applications. One way of extracting these arguments, their semantic classes, and their combinations is by corpus analysis.

In our study of *pollution,* the corpus used for the extraction of linguistic information was the EcoLexicon English Corpus (over 54,000,000 words), which was subsequently validated by the English TenTen corpus (EnTenTen) of Internet texts, compiled by Lexical Computing. This English corpus contains over 19 billion words and is tagged with TreeTagger using the UTF-8 parameter file. The linguistic information was automatically extracted with the Sketch Engine application (www.sketchengine.eu). One of its most useful functionalities is the word sketch, which is an automatic corpus-derived summary of a word's grammatical and collocational behavior (Kilgarriff et al. 2014).

Based on the corpus information extracted from concordances of *pollute* and its different forms, Table 2 shows that the most frequent polluting agents or contaminants belong to the semantic categories of HUMAN ACTIVITY, INDUSTRY, WASTE, CHEMICAL, GAS EMISSION, VEHICLE, and MICROORGANISM.

In contrast, the second argument, which is the polluted entity, consists of different specifications of AIR, WATER, and SOIL.

Table 2. Semantic classes of the arguments of *pollute*

ARG 1 **Polluting Agent**	**Contaminant**
Human activity	[**Activity**] fracking, drilling, mining
Industrial location	[**Location**] factory, power plant, mine
Waste	[**Solid**] garbage, landfill, sludge [**Liquid**] effluent, wastewater, runoff
Chemical	[**Element**] mercury, carbon, nitrogen, phosphorus [**Natural mixture**] coal, oil, petroleum [**Artificial mixture**] pesticide, fertilizer
Gaseous emission	**[Industrial source**] gases, fumes

ARG 1 **Polluting Agent**	**Contaminant**
	[**Vehicle source**] exhaust
Vehicle	[**Land vehicle]** car, diesel vehicle [**Water vehicle**] container ship, oil tanker [**Air vehicle**] aircraft, jet
Microorganism	bacteria
POLLUTES	
ARG 2 **Polluted Entity**	**Environmental element/location**
Environment	environment
Water	[**Water**] water, groundwater, drinking water [**Water body**] aquifer, river, ocean, stream, creek, watershed, lake
Air	[**Gas**] air, airwaves, atmosphere
Soil	[**Soil**] land, soil, ground, Earth

What is important is not the syntactic realization of the predicate and its nominalization, but rather the combination of semantic roles and categories, which reflect the polluting activities of the human race (since the implicit agent is human) as well as the three main environmental spheres (air, water, and soil) where pollution occurs. Consequently, the frame is generated by this combination of semantic roles and categories, in this case, POLLUTING AGENT (CONTAMINANT) and POLLUTED ENTITY (ENVIRONMENTAL ELEMENT/LOCATION) and the relation between them.

3.3. From phrase level to frame level

The understanding of phrases in specialized language depends on the reader's ability to expand them so that they fit within a wider context or frame. The problem is that frames are slippery customers. Everyone talks about them but examples are rarely provided, except for the much-used example of the commercial transaction (Fillmore 1982). However, frames also exist in specialized language and can be specified for the knowledge domains, such as the environment (Faber 2012, 2015).

Generally speaking, a frame is a type of mental representation, involving the organization of knowledge about a concept or a set of related concepts. The elements within a frame are linked by different types of semantic relation (Minsky 1975; Fillmore 1985, 2006; Faber 2012, 2015).

The specification of a specialized language frame is the description of a space and the events that occur within it as well as the entities that participate in those events. Busse (2012) makes the useful distinction between concept frames and

predicative frames. Concept frames mostly refer to concepts designated by nouns and noun phrases. Concept frames represent the attributes and properties of an entity. As such, they provide a general format for the representation of categories and category structure (Barsalou 1992). In contrast, predicative frames describe actions and processes, which are designated by verbs and their nominalizations. They represent events and states of affairs in terms of their situation types and participants.

Evidently, predicative frames are more relational since they are composed of various concepts. For this reason, they are the most useful for text understanding. They not only arise from single verbs but also from general configurations of verb meaning that converge in a single semantic space. In specialized language, this sounds strange because verbs are rarely regarded as terms, and thus usually not included in specialized knowledge resources (L'Homme 1998; Buendía Castro 2013). However, general language verbs are crucial to meaning because they are generally what relate concepts in specialized texts.

For example, of the 703 most frequent verbs in the EcoLexicon corpus of over 54 million words, only 10 verbs have no general language meaning (*denitrify, flocculate, hybridize, mineralize, nucleate, oxygenate, photosynthesize, solubilize, subduct,* and *supercool*). The other verbs are general language verbs (e.g. *accumulate, increase, develop, produce, vanish, pollute* etc.), which are also used in specialized texts with terms as their arguments. Their meaning underlies what happens in the environment and how we talk about it.

Even though verbs (especially general language verbs) have never been regarded as important in Terminology, they reflect how environmental entities interact. These verbs represent what in our opinion are conceptual invariants, which are present in the majority of documented language-cultures. The existence of such unique beginners or semantic near primitives that are lexicalized in most languages is a constant in the work of linguists such as Ana Wierzbicka, George Miller, and Juri Apresjan, *inter alia*. This culturally shared knowledge, stored in the lexicon, is composed of stable points of reference that comprise a cognitive map of our phenomenological universe.

In previous research within the framework of the Lexical Grammar Model, Faber and Mairal (1999) analyzed and categorized the semantic and syntactic structure of 12,000 general language verbs, first in English and subsequently in Spanish. This resulted in the following general lexical domains: EXISTENCE (*be, happen*), CHANGE (*become, change*), POSSESSION (*have*), SPEECH (*say, talk*), EMOTION (*feel*), ACTION (*do, make*), COGNITION (*know, think*), MOVEMENT (*move, go, come*), PHYSICAL PERCEPTION (*see, hear, taste, smell, touch*), MANIPULATION (*use*), CONTACT/IMPACT (*hit, break*) and POSITION (*to put, to be*). Other classes included LIGHT, SOUND, BODY FUNCTIONS, WEATHER, etc.

Faber and Mairal (1999) used this inventory of verb classes to classify the most general environment-related actions and processes in lexical domains derived from definition factorization, as described in the Lexical Grammar Model. This

highlighted the most prominent actions and processes within the environment as well as the semantic categories of the typical participants in these event frames.

For example, when the 703 most frequent verbs in the EcoLexicon corpus were analyzed, the majority were found to belong to the lexical domains, dimensions, and subdimensions of CHANGE, MOVEMENT, EXISTENCE, POSSESSION, POSITION, IMPACT, and MANIPULATION. Table 3 shows some of the verbs that belong to these lexical domains.

Table 3. Organization of verbs in lexical domains

Lexical domain	Verb examples
CHANGE [to become/change]	*abate, accrete, aggravate, ameliorate, clarify, decrease, deform, enrich, pollute, contaminate,* etc.
MOVEMENT [to move]	*whirl, vibrate, topple, thrust, submerge, spiral, stir, shake rotate*, etc.
EXISTENCE [to be/exist]	*prevent, produce, obliterate, originate, occur, interrupt, initiate, inhibit, inhabit*, etc.
POSSESSION [to have]	*absorb, catch, cling, collect, conserve, cumulate, drain, entrap, exchange, grab, grasp, harvest*, etc.
POSITION [to be in a state/place/position]	*block, cover, dump, embed, encase, encrust, envelop, juxtapose, lodge, plug, replace*, etc.
MANIPULATION [to use]	*burn, consume, exert, expend, exhaust, ignite, irradiate, recycle, tape, use, utilize,* etc.
IMPACT [to hit/break]	*thresh, strike, slam, shatter, rupture,* etc.

Notably absent was the (frequent) use of verbs belonging to the areas of FEELING, SENSORY PERCEPTION, and SPEECH. What is even more interesting is that the verbs in the same lexical domain tended to combine with specialized knowledge units in the same or similar semantic classes such as LIQUID SUBSTANCE, SOLID SUBSTANCE, CHEMICAL ELEMENT, WEATHER EVENT, LANDFORM, WATERBODY, etc.

Faber and Mairal (1999) highlighted the fact that one of the most important environmental processes is CHANGE. CHANGE is a lexical domain with a number of dimensions, which are specific to variation in parameters of time, space, and evaluation (e.g. *to become better, to become worse, to become larger, to become smaller,* etc.). *Pollute*, for example, belongs to this lexical domain. A segment of the lexical domain of CHANGE (*To cause something to become worse*) is shown in Table 4.

Table 4. Segment of the lexical dimension *To cause sth to become worse* in the lexical domain of **CHANGE**

LEXICAL DOMAIN OF CHANGE TO CAUSE STH TO BECOME WORSE (IMPURE/DANGEROUS/UNCLEAN)
contaminate to cause sb/sth to become worse by making it less pure.
pollute to contaminate sth (esp. water/air/soil), by adding a harmful substance to it so that it is dangerous to the health of living organisms.
adulterate to pollute sth (esp. food products) by adding sth that lowers quality esp. to defraud the user.
taint to pollute (sth) so that it is spoiled or damaged.
poison to contaminate sb/sth by adding a harmful substance to it so that it will die or make others die.
infect to contaminate sb/sth with disease-producing organisms.

As can be observed, *pollute, poison,* and *infect* are hyponyms of *contaminate,* which is the most general term in this subdimension. The difference between *pollute, poison,* and *infect,* lies in the polluting substance or what is polluted. When the semantic (and syntactic) characteristics of the verbs are also specified, this type of lexical organization codifies the range of choices available to each speaker in the lexicalization of a given area of meaning.

The assumption here is that verbs within the same lexical subdimension have a similar syntax and, even more important, combine with the same semantic types of argument. In the case of specialized language, the polysemy of these general language verbs is limited because the scope of their meaning is restricted to the field of Environmental Science. However, verb meaning is not restricted by syntax, but rather the nature of their arguments, which belong to a set of specific conceptual categories such as LANDFORM, CHEMICAL ELEMENT, ATMOSPHERIC PHENOMENON, WATER BODY, PLANT, etc.

The POLLUTION frame can also be further extended to include verbs that codify the remedy for pollution, in this case, a cleaning action in the form of the polysemic general language verb *flush.* Depending on whether there is a focus on liquid movement (*flowing*) or the result (*cleansing*), it is a member of the lexical domain of MOVEMENT or CHANGE. Although *flush* is polysemic, it only has one meaning in Environmental Science. The semantic nature of its arguments is what restricts its meaning to movement in a liquid medium. Its definition is the following:

(2)

flush to cause a **liquid** to flow into/through [MOVEMENT] a **place**, cleaning it [CHANGE] of **something**.

It thus activates a frame with three arguments or participants: (i) a liquid; (ii) a place; and (iii) an (undesirable) substance. In the EcoLexicon corpus, these argument slots are filled by the following terms in the following semantic classes, as illustrated in Table 5.

Table 5. Terms and semantic classes that can fill the argument slots of *flush*

<table>
<tr><td colspan="3">flush to cause a <u>liquid (usu. water)</u> to flow into/through a <u>place</u>, cleaning it of <u>something</u></td></tr>
<tr><th>Argument 1: Liquid</th><th>Argument 2: Place</th><th>Argument 3: Substance</th></tr>
<tr><td>WEATHER EVENT → storms/rainfall</td><td rowspan="3">WATER BODY→
ENCLOSED → lagoon, pond, lake
SEMI-ENCLOSED → estuary, harbor, basin, embayment, river, bay
OPEN → beach, channel, slope, reef</td><td>SOLID SUBSTANCE → sediment, sand, littoral material</td></tr>
<tr><td>WATER → water
COMPOSITION → freshwater, saltwater, salt brine, seawater
VELOCITY → water cascades
QUANTITY→ flood</td><td>LIQUID SUBSTANCE→ soil water, acid, dissolved metals
CHEMICAL ELEMENT → magnesium, sodium</td></tr>
<tr><td>SEA/OCEAN MOVEMENT→ tides, tidal currents, tidal action</td><td>HARMFUL SUBSTANCE → pollutant, contaminant, organic matter, harmful salts, acid, dissolved metals</td></tr>
</table>

As can be observed, each definitional slot is potentially filled by a specific set of semantic types and subtypes. In this sense, each argument generates a mini-ontology. The frame activated pertains to water movement into a water body, resulting in the cleansing of that place of a usually harmful substance. In this sense, *flush* is a predicate that is related to pollution, and which provides a subframe that relates clusters of semantic categories that represent entities in the environment.

Evidently, this type of context specification enhances understanding since it identifies the types of entity that participate in events. The focus here is on the actions and processes designated by verbs. As previously mentioned, when specialized knowledge units fill their respective argument slots, the meaning of these general language verbs is constrained by the semantic categories of their arguments. This highlights the relational potential of predicative frames and their usefulness for specialized knowledge acquisition.

4. Conclusion

In this paper we have described how knowledge acquisition can be conceived as a progressive expansion of meaning, which begins at the term-level, progresses to the phrase level, and finally results in the codification of an entire knowledge

frame. In this sense, the definition of a single term can predict how its meaning can be potentially specified in MWTs (Cabezas-García and Faber in press). MWTs that designate processes can be represented in terms of their predicate-argument structure. The importance of exploring semantic types and their combinations cannot be overstressed because semantics, rather than syntax, is what can disambiguate MWTs and phrases in specialized texts (Buendía Castro and Faber 2016). This was also evident in the analysis at the phrase and frame levels, where semantic categories and roles were found to be the basis for knowledge activation.

As reflected in our analysis of *pollution*, it is also necessary to take a closer look at the semantics of general language verbs in specialized texts. They show how specialized knowledge units are combined and encode the basic activities, processes and events in a specialized domain. The specification of context is a way of clarifying the meaning of the terms in a text. The examples given highlight the usefulness of using language as a conceptual mirror that reflects how specialized knowledge is structured and configured.

References

Anderson, John R. 1983. *The Architecture of Cognition*. Hillsdale: Lawrence Erlbaum.

Binder, Jeffrey R. and Rutvik H. Desai. 2011. The neurobiology of semantic memory. *Trends in Cognitive Sciences* 15 (11). 527-536.
[Online]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3350748/.
[Accessed: 22 April 2018].

Buendía Castro, Miriam. 2013. *Phraseology in Specialized Language and its Representation in Environmental Knowledge Resources*. PhD Thesis. Granada: University of Granada.

Buendía Castro, Miriam and Pamela Faber. 2016. Phraseological Correspondence in English and Spanish Specialized Texts. In Gloria Corpas Pastor (ed.), *Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives = Fraseología computacional y basada en corpus: perspectivas monolingües y multilingües*, 391-398. Geneva: Tradulex.

Busse, Dietrich. 2012. *Frame-Semantik: Ein Kompendium*. Berlin and Boston: De Gruyter.

Cabezas-García, Melania and Pamela Faber. In press. Phraseology in specialized resources: An approach to complex nominals. *Lexicograpy*.

Cabezas-García, Melania and Pilar León-Araúz. 2018. Towards the Inference of Semantic Relations in Complex Nominals: a Pilot Study. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, 2511-2518. Miyazaki: ELRA.

Daille, Béatrice. 2017. *Term Variation in Specialized Corpora: Characterisation, automatic discovery and applications*. Terminology and Lexicography Research and Practice, vol. 19. Amsterdam and Philadelphia: John Benjamins.

Dubuc, Robert and Andy Lauriston. 1997. Terms and contexts. In Sue Ellen Wright and Gerhard Budin (eds.), *Handbook of Terminology Management: Basic Aspects of Terminology Management*, vol. 1, 80-87. Amsterdam and Philadelphia: John Benjamins.

Faber, Pamela. (ed.) 2012. *A Cognitive Linguistics View of Terminology and Specialized Language*. Berlin and New York: Mouton de Gruyter.

Faber, Pamela. 2015. Frames as a framework for terminology. In Hendrik J. Kockaert and Frieda Steurs (eds.), *Handbook of Terminology*, vol. 1, 14-33. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Faber, Pamela and Pilar León-Araúz. 2016. Specialized knowledge representation and the parameterization of context. *Frontiers in Psychology* 7 (196). 1-20. [Online] Available from: https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00196/full. [Accessed: 19 April 2018].

Faber, Pamela and Ricardo Mairal Usón. 1999. *Constructing a Lexicon of English Verbs*. Berlin: Mouton de Gruyter.

Faber, Pamela, et al. 2014. Neural Substrates of Specialized Knowledge Representation: An fMRI study. *Revue française de linguistique appliquée* 19 (1). 15-32. [Online] Available from: https://www.cairn.info/revue-francaise-de-linguistique-appliquee-2014-1-page-15.htm. [Accessed: 20 April 2018].

Fernández-Domínguez, Jesús. 2016. A morphosemantic investigation of term formation processes in English and Spanish. *Languages in Contrast* 16 (1). 54-83.

Fillmore, Charles J. 1982. Frame Semantics. In The Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*, 111-137. Seoul: Hanshin.

Fillmore, Charles J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* 6 (2). 222-254.

Fillmore, Charles J. 2006. Frame Semantics. In Dirk Geeraerts (ed.), *Cognitive Linguistics. Basic readings*, 373-400. Berlin and Boston: De Gruyter.

Fillmore, Charles J., Johnson, Christopher R. and Miriam R. L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography* 16 (3). 235-250.

Gallese, Vittorio and George Lakoff. 2005. The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology* 22 (3-4). 455-479.

Gildea, Daniel and Daniel Jurafsky. 2002. Automatic Labeling of Semantic Roles. *Computational Linguistics* 28 (3). 245-288. [Online] Available from: https://dl.acm.org/citation.cfm?id=643093. [Accessed: 22 April 2018].

Kecskes, Istvan. 2014. *Intercultural Pragmatics*. Oxford and New York: Oxford University Press.

Kiefer, Markus and Friedemann Pulvermüller. 2012. Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex* 48 (7). 805-825.

Kilgarriff, Adam et al. 2014. The Sketch Engine: ten years on. *Lexicography* 1 (1). 7-36.

Levi, Judith. 1978. *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.

Levin, Beth. 2013. Argument Structure. In Mark Aronoff (ed.), *Oxford Bibliographies in Linguistics*. New York: Oxford University Press.

L'Homme, Marie-Claude. 1998. Le statut du verbe en langue de spécialité et sa description lexicographique. *Cahiers de lexicologie* 73 (2). 61-84.

Lyons, John. 1995. *Linguistic semantics: an introduction*. Cambridge and New York: Cambridge University Press.

Maniez, François. 2009. L'adjectif dénominal en langue de spécialité: étude du domaine de la médecine. *Revue française de linguistique appliquée* 14 (2). 117-130. [Online] Available from: https://www.cairn.info/revue-revue-francaise-de-linguistique-appliquee-2009-2-page-117.htm. [Accessed: 19 April 2018].

Meteyard, Lotte et al. 2012. Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48 (7). 788-804.

Minsky, Marvin. 1975. A framework for representing knowledge. In Patrick H. Winston (ed.), *The Psychology of Computer Vision*, 211-277. New York: McGraw-Hill.

Nakov, Preslav. 2013. On the interpretation of noun compounds: Syntax, semantics, and entailment. *Natural Language Engineering* 19 (3). 291-330. [Online] Available from: https://pdfs.semanticscholar.org/08af/6f1481ca99a07f01b5dfde042ea00d1bf8b5.pdf. [Accessed: 20 April 2018].

Palmer, Martha, Gildea, Daniel and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics* 31 (1). 71-106. [Online] Available from: https://dl.acm.org/citation.cfm?id=1122628. [Accessed: 20 April 2018].

Patterson, Karalyn, Nestor, Peter J. and Timothy T. Rogers. 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience* 8. 976-987.

Quillian, Ross. 1969. The teachable language comprehender. *Communications of the ACM* 12 (8). 459-476.

Sager, Juan C., Dungworth, David and Peter F. McDonald. 1980. *English Special Languages. Principles and Practice in Science and Technology*. Wiesbaden: Brandstetter Verlag.

Štekauer, Pavol, Valera, Salvador and Lívia Körtvélyessy. 2012. *Word-formation in the world's languages: a typological survey*. Cambridge: Cambridge University Press.

Temmerman, Rita. 2013. Primary and secondary term creation and the process of understanding. Presentation given at the international seminar *Application of Cognitive Terminology Theories in Terminology Management*, Zagreb, 27-28 September 2013. [Online]. Available at: http://struna.ihjj.hr/uploads/actt/Temmerman_Zagreb_2013.pdf. [Accessed: 17 April 2018].

Van Valin, Robert D. and Randy J. LaPolla. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.