

Title Page

Lorella Viola

School of Politics, Philosophy and Languages and Communication Studies, University of East
Anglia Norwich Research Park NR4 7TJ Norwich, United Kingdom

Phone: +44 (0) 1603 593411, Mobile: +44 (0) 7449336636, E-mail: L.Viola@uea.ac.uk

A corpus-based investigation of language change in Italian: The case of *grazie/ringraziare di* and *grazie/ringraziare per*

Abstract

In Italian, *grazie* ‘thanks’ and *ringraziare* ‘to thank’ historically introduce a recipient by means of the preposition *di* ‘of’ (Renzi et al 1991: 545-548); when *grazie* and *ringraziare* introduce a subordinate infinite clause, they may all the same be followed by either *di* or *per* ‘for’, the latter being the habitual preposition introducing an implicit causal subordinate (Renzi et al *ibid.*). In light of these considerations, a general lower frequency of occurrence of collocations with *per* would be expected. However, a number of authors (e.g., Renzi 2000, Alfieri et al 2008: 331) have reported an increase in the use of constructions with *per*; though with differences in the approach and the framework employed, they have also hypothesised that such an increase may be due to language contact with English. A careful exploration of the relevant literature, however, has revealed that such claims of both an increase in the use of *grazie/ringraziare per* in Italian and of an influence from English as the cause of the increase have so far outpaced empirical substantiation. This study, on the contrary, uses verifiable and objective data such as diachronic language corpora of written, spoken and dubbed Italian to empirically investigate the distribution of both constructions through the history of Italian. The results will reveal that, from 1200 to 2011, the frequency of use of forms with *per* has indeed more than octupled in writing and that, from 1965 to 2004, has more than doubled in speech. Moreover, by analysing the distribution of the studied constructions in a corpus of dubbed Italian from (American) English, the article will also explore the possibility that language contact with English, mainly via dubbing translations, may have played a concurrent fundamental role for such changes.

1. Introduction

This article investigates the diachronic trend of two Italian constructions, *grazie/ringraziare di* ‘thanks/to thank of’ (thanks/to thank for) and *grazie/ringraziare per* ‘thanks/to thank for’. It is argued that over time *grazie/ringraziare per* have increasingly been used in Italian over *grazie/ringraziare di*, the latter considered the historically preferred form (Renzi et al 1991: 545-548). The claims of the increase in the use of constructions with *per* has been reported by a number of authors (e.g., Renzi 2000, Alfieri et al 2008: 331) who, though with differences in the way they approached the subject, have also hypothesised that such an alleged increase may be due to an influence from the English *thanks/to thank for*. Renzi, for example, within the supposed increase of forms with *per*, mainly focusses on the specific use of *grazie/ringraziare per* when thanking

someone in anticipation of future situations; this use - not admissible in contemporary Italian (cfr. Serianni 2000) – would be reflected in sentences such as *grazie di/per mandare la lettera* (thanks for sending the letter). According to the scholar, the appearance in Italian, particularly in business contexts, of constructions where *grazie/ringraziare per* are used to refer to future actions would be an indication of the interference from English. However, the alleged increase in such constructions in Italian, whether referred to future situations or not, is not supported by any verifiable data and solid evidence is overall missing.

Similarly, Alfieri et al (2008) argue that there has been an increase in constructions followed by the preposition *per* and that it may be due to the influence from English; her hypothesis, however, is that the main source of interference would be the dubbing of (American) English films and TV programmes. Such a hypothesis lies upon the fact that the Italian language is, arguably, particularly subject to this type of interference as in Italy, over 90% of all the audiovisual (AV) products are imported (Cinetel 2016), and therefore dubbed; moreover, dubbing has been steadily in use since 1932. However, their claims of changes in the Italian language because of an influence from English via dubbing have so far only been limited to descriptive approaches, i.e., with no or negligible use of empirical data and a systematic investigation of the extent to which such instances of interference may have passed into real use Italian has not been conducted yet.

Specifically, these authors do make use of quantifiable data such as corpora of both dubbed and original Italian TV programmes to substantiate their claims, but there are limitations in both their analyses and the way they present their results. For example, the corpora themselves are not available and only partial and nonspecific details of the frequency of occurrence of *grazie/ringraziare di* vs *grazie/ringraziare per* (raw or relative) are given. The size of the corpora is indicated in hours of recorded TV programmes rather than per number of words, therefore the terms of comparison used to qualify the given frequencies of occurrence as high are not clear; furthermore, the claims of why uses of *grazie/ringraziare di* are considered ‘traditional’ (Alfieri et al 2008: 331) are not supported by any lexicographic or diachronic quantitative investigation. In other words, though occurrences of *grazie/ringraziare per* were found higher than those with *di* in both dubbed products from (American) English and original Italian TV programmes, their results cannot be considered as conclusive.

On the whole then, it appears that regardless of the approach and the framework employed, both the claims of an increase in real use Italian of *grazie/ringraziare per* and the alleged influence from English have been hypothesised rather than effectively proven. On the contrary, the aim of this study is to use empirical data, such as language corpora of written, spoken and dubbed Italian, to investigate the diachronic distribution of *grazie/ringraziare di* and *grazie/ringraziare per* in the

language so as to obtain a full account of the trend of these constructions through the history of Italian. By analysing the distribution of the studied constructions over time, particularly before and after the coming of dubbing in Italy as well as in Italian dubbed products from (American) English, it will also be possible to explore the research hypothesis that an influence from English via dubbing may be claimed. In this respect, it is important to clarify that the study does not try to claim that language change can occur without live social interaction (Giles & Powesland 1975, Giles 1984, Trudgill 1986 Giles et al 1991, Milroy 1992, 2002, Labov 1994, 2001, 2010, Eckert 2000, 2008) which clearly plays a fundamental role in diffusing certain language features, nor does it claim that dubbing is the only direct cause for the increase in the use of the forms under analysis. For instance, factors such as globalisation, the Internet, and the predominance of English as lingua franca are acknowledged as other major sources of influence. The investigations are therefore carried out to empirically assess 1) if *grazie/ringraziare per* constructions are historically less frequent than *grazie/ringraziare di* forms, 2) the frequency of use of both constructions in Italian dubbed products from (American) English and 3) if and when the distribution of the two forms has changed thus helping shed the light on plausible correlations with the coming of dubbing.

First, I will outline the methodology adopted in the study and provide a brief overview of the resources used for the investigation (§ 1.2). In § 2, I will then present the data showing the distribution of the two forms through the history of Italian and in a corpus of Italian films dubbed from (American) English which has been built for the purposes of this study. Conclusions are finally drawn in § 3.

1.2 Methodology and resources

The research hypothesis of this study is that, over time, *grazie* and *ringraziare* have been increasingly used in combination with *per* over forms followed by *di* which are considered to be traditionally more used (Renzi et al 1991: 545-548). The analysis in this way needs to account not only for the frequency of occurrence of the two constructions in old Italian, but crucially, also for their distribution through the history of Italian. This is paramount to assess if it is true that traditionally *grazie/ringraziare* collocate more strongly with *di* than with *per* so as to conclusively establish whether an increase in collocations with *per* has occurred. If evidence is gathered that such a change has taken place, then the article will also explore the hypothesis that the increase may be correlated to an influence from the English *thanks/to thank for* via dubbing translations. Thus, the distribution of the two forms will be investigated in our corpus of Italian dubbed films from (American) English so as to identify any possible significant differences between the frequency of occurrence of the two constructions; the results will be then compared with the data of their

distribution in real use Italian. The procedure will allow us to identify possible positive correlations with Italian dubbing tradition thus ultimately validating the hypothesis of a plausible influence from English via dubbing.

Unlike previous accounts, this study develops upon an evidence-based approach where in-depth rigorous investigations are pursued across a range of verifiable data such as etymological and historical dictionaries together with corpora of old, contemporary (both written and spoken) and dubbed Italian. The written corpora of old Italian used for the analysis are the OVI Corpus (Opera del Vocabolario Italiano) which gathers texts of written Italian up to 1375 (about 23 millions of words) and the MIDIA corpus (Morfologia dell'italiano in Diacronia) which collects texts from the thirteenth century to the early twentieth century (about 7,5 million words). The written corpora of modern and contemporary Italian (DiaCORIS and CORIS) gather authentic Italian texts from 1861 to 2011 (about 160 million words) while the spoken corpora (Stammerjohann, LIP, C-ORAL-ROM) collect oral dialogues from 1965 to 2003 (about 1 million words). To explore the research hypothesis that an increase in *grazie/ringraziare per* constructions may be due to the influence of English, particularly via dubbing translations, a corpus of Italian films dubbed from (American) English will also be used; the corpus has been built by collecting the scripts of 15 dubbed films¹ distributed in Italy between 1964 and 2007 (about 155 thousand words). The list of resources is not to be considered complete or finished by any means and in future works, different resources may be used, either as complementary or substitute tools for those employed here; the complete list of the resources employed here is provided in the reference list.

These linguistic resources have been selected according to criteria of completeness, authoritativeness, and representativeness of diaphasic, diastratic, diamesic, diatopic, and diachronic variation and sample a broad range of authors and genres which may be considered to even out and provide a reasonably accurate picture of written Italian as a whole and of cities, speakers and contexts for spoken Italian. As argued by McEnery & Wilson (2001: 78), the criticism that frequency rates may be unrepresentative of the population as a whole, for example when they are particularly low, applies “not only to linguistic corpora but to any form of scientific investigation which is based on sampling rather than on the exhaustive analysis of an entire and finite population.” However, because the corpus is sampled to be maximally representative of the population, findings on that sample may be generalised to the larger population and furthermore, it means that direct comparisons may be made between different corpora. Conversely, when there are no occurrences, this is also an interesting and important comment on the frequency of that specific construct or word (ibid.). There will always be the possibility that some constructions may occur due to pure chance, but such limitations – which again apply to any sampling analysis - can at least,

in corpus linguistics, be partially addressed by maximising representativeness. Furthermore, significance tests (e.g., chi square value, df value, *p*-value, log likelihood test) will be performed to exclude the possibility that any observed effect will have occurred due to a sampling error alone (cfr. Babbie 2013).

Finally, because the corpora used are different in size, the results of the quantitative investigations are presented in tables which show the number of occurrences in each corpus (raw frequency) and the corresponding proportions in parts per million (p.p.m.) for the whole corpus and for each time period, when applicable.

2. *Distribution of 'grazie/ringraziare di' and 'grazie/ringraziare per' through the history of Italian*

This section investigates the diachronic distribution of *grazie/ringraziare di* and *grazie/ringraziare per* through the history of Italian in order to assess which form was historically preferred and if any change has happened over time. The detailed account of their frequency of occurrence and diachronic trends will provide us with valuable information which will be relevant to the research hypotheses of both an increase in the use of forms followed by *per* and any possible plausible correlation with the influence of English from dubbing. Section 2.1 investigates the distribution of the two variants in Old Italian (up to 1375), section 2.2 analyses the respective frequencies of occurrence from the 13th century to the 20th century. In § 2.3, the investigations are carried out to cover the period from 1861 to 2011; while the distribution in dubbed films from 1965 to 2007 is analysed in § 2.4, § 2.5 explores both forms in spoken Italian.

2.1 *Distribution in Old Italian (up to 1375)*

The Italian *grazie* 'thanks' is the elliptic form of *vi rendo grazie* 'I will return you the favour'; as for the verbal phrase it derives from, *grazie* introduces a beneficiary by means of *di* 'of', which is considered to be the preposition historically preferred (Renzi et al 1991: 545-548). However, when *grazie* and *ringraziare* introduce a subordinate infinite clause, they may be followed by either *di* or *per* 'for', the latter being the habitual preposition introducing an implicit causal subordinate (ibid.). The etymological (DELI) and lexicographic (Crusca 1863-1923, vol. VII: 565) sources report occurrences for *ringraziare di* but no occurrences for constructions with *per* have been found thus suggesting that the preposition *di* historically collocates more strongly than *per*. In this section, the distribution in Old Italian of *grazie/ringraziare di* and *grazie/ringraziare per* is investigated in order to further verify the etymological and lexicographic findings. The corpus used for the analysis is the OVI Corpus which collects texts up to 1375; the results are shown in tables 1 and 2 below.

Table 1: OVI - Frequency rates of *grazie di* and *grazie per*

OVI	GRAZIE DI	GRAZIE PER
UP TO 1375	23	7
P.P.M.	1	0.3

Table 2: OVI - Frequency rates of *ringraziare di* vs *ringraziare per*

OVI	RINGRAZIARE DI	RINGRAZIARE PER
UP TO 1375	106	15
P.P.M.	4.61	0.65

The results show that up to 1375, *grazie/ringraziare di* collocated more strongly than *grazie/ringraziare per* thus confirming that historically, forms followed by *di* were preferred in Italian. Here below, [1], [2], [3] and [4] are examples retrieved from the corpus of the contexts of use of the constructions under analysis (bold mine).

[1] Io re di Francia faccio molte **grazie delle** grandi proferte (E/KI) (1282-99)

I, King of France, give thanks for the great offers

[2] In comandamento abbiamo di rendere **grazie per** queste cose (Cit/1. Deche) (14th century)

We have been commanded to give thanks for these things

[3] Direte al nostro Signore lo **ringraziamo di** tanta buona proferta (E/KI) (1282-99)

You will tell to our Lord that we will thank him for such good offer

[4] **Per** queste parole, **ringrazia** l'altore Beatrice (Cit/1. Chiose) (1375)

For these words, Beatrice thanks he who gives life

In the examples above it can be noticed that there are no observable differences in the contexts of use, conveyed meaning or function of the two forms thus showing that the prepositions can be used interchangeably. In the next section, I will investigate the frequency of occurrence of both forms in the MIDIA Corpus which collects data from the thirteenth century to the first half of the twentieth century.

2.2 Distribution from 13th century to 20th century

I will now move on to analyse the distribution of *grazie/ringraziare di* and *grazie/ringraziare per* from 1200 to 1947. Tables 3 and 4 show the raw frequency rates and the correspondent p.p.m.

proportions for occurrences of both constructions; the relative frequencies are calculated both over the whole corpus (7,652,526 words) and for each time slot, which is divided as it follows: 1200-1375 (1,238,457 words); 1376-1532 (1,646,428 words); 1533-1691 (1,600,301 words); 1692-1840 (1,499,412 words); 1841-1947 (1,667,928 words). This will allow us to compare not only potential distribution differences of the frequencies over the whole time period covered by the corpus, but also within the same time slot. Finally, significance tests (e.g., chi-square value, df value, p-value) will be performed to exclude the possibility that any observed effect will have occurred due to a sampling error alone (cfr. Babbie 2013).

Table 3: MIDIA - Frequency rates of *grazie di* and *grazie per*

MIDIA	GRAZIE DI	P.P.M.	P.P.M. (TOT.)	GRAZIE PER	P.P.M.	P.P.M. (TOT.)
1200-1375	2	1.61	0.26	1	0.81	0.13
1376-1532	7	4.25	0.91	2	1.21	0.26
1533-1691	10	6.24	1.30	4	2.50	0.52
1692-1840	8	5.33	1.04	4	2.67	0.52
1841-1947	19	11.39	2.48	2	1.19	0.26
TOT.	46	/	5.99	13	/	1.69

Table 4: MIDIA - Frequency rates of *ringraziare di* and *ringraziare per*

MIDIA	RINGRAZIARE DI	P.P.M.	P.P.M. (TOT.)	RINGRAZIARE PER	P.P.M.	P.P.M. (TOT.)
1200-1375	5	4.04	0.65	2	1.61	0.26
1376-1532	18	10.93	2.35	2	1.21	0.26
1533-1691	23	14.37	3	0	0	0
1692-1840	48	32.01	5.62	8	5.33	1.04
1841-1947	30	17.99	3.92	16	9.59	2.09
TOT.	124	/	16.2	28	/	3.66

The results show that through the history of Italian, *grazie/ringraziare di* collocates more strongly than *grazie/ringraziare per*; also, a direct comparison between the different time periods gives clear evidence of that collocations with *di* are consistently more frequent than collocations with *per*. In order to establish whether the difference in the forms' frequencies is significant,

significance tests such as chi square value, df value, p-value have been performed. The chi-square value result for *grazie di* and *grazie per* is 3.38, df value is 4 and the *p-value* is 0.4969128 while for *ringraziare di* and *ringraziare per* the chi square value is 15,45, df value is 4 and the *p-value* is 0.003855315. These results show that the difference of frequency is significant for *ringraziare di* and *ringraziare per*; although *grazie di* occurs more frequently than *grazie per*, the difference in their frequency is not significant. Such a result will be relevant when compared with the difference in their frequency distributions in more recent corpora; if found significant, it will indeed evidence that a substantial change has occurred.

The quantitative analysis has then revealed that constructions with *di* consistently collocate more strongly than forms with *per*; a qualitative analysis of the occurrences is now performed (examples below [5], [6], [7] and [8] below) to identify any potential observable difference in the contexts of use of the two constructions (bold mine).

[5] Insieme meco **grazie** a Dio rendete **dell'**ammirabili sua piet  divina (TEA2_LMED_RAPGP) (1376-1532)

Join me in thanking God for his admirable divine compassion

[6] Ella rendea cortese **grazie per** lodi (POE3_TAS_GERU00) (1533-1691)

She kindly thanked for the prayers

[7] Sperate in Dio, seguendo suo dottrina, **ringraziandol d'**ogni beneficio (TEA2_RAPP_GRLAZ) (1376-1532)

Have hope in God, follow his doctrine, and thank him for his help

[8] Leggiamo di fatti [...] che S. Gregorio lo **ringrazia per** i quattrocento scudi d'oro (PER5_DESA_ROM00) (1841-1947)

We read that Saint Gregorio thanks him for four hundred gold ecus

As no significant difference in the contexts of use, conveyed meaning or function can be observed, the examples above confirm that the two constructions are equivalent. These results are consistent with the previous findings and confirm that *grazie/ringraziare di* is the form historically preferred. In the next section, I will investigate the frequency of occurrence of both forms in the DiaCORIS and CORIS Corpus which collect data from the 1861 to 2011.

2.3 Distribution from 1861 to 2011

This section presents the results of the investigation of *grazie/ringraziare di* and *grazie/ringraziare per* in the DiaCORIS and the CORIS Corpus. Tables 5, 6, 7 and 8 show the raw

frequency rates and the correspondent p.p.m. proportions for occurrences of both constructions; the relative frequencies are calculated both over the whole corpus (25 million words) and for each time slot, which approximately contains 5 million words.

Table 5: DiaCORIS - Frequency rates of *grazie di* and *grazie per*

DiaCORIS	GRAZIE DI	P.P.M.	P.P.M. (TOT.)	GRAZIE PER	P.P.M.	P.P.M. (TOT.)
1861-1900	12	2.4	0.48	0	0	0
1901-1922	17	3.4	0.68	0	0	0
1923-1945	4	0.8	0.16	0	0	0
1946-1967	13	2.6	0.52	3	0.6	0.12
1968-2001	3	0.6	0.12	3	0.6	0.12
TOT.	49	/	1.96	6	/	0.24

Table 4: CORIS - Frequency rates of *grazie di* and *grazie per*

CORIS	GRAZIE DI	P.P.M.	GRAZIE PER	P.P.M.
1980-2011	290	2.23	485	4.72

Table 7: DIACORIS - Frequency rates of *ringraziare di* and *ringraziare per*

DiaCORIS	RINGRAZIARE DI	P.P.M.	P.P.M. (TOT.)	RINGRAZIARE PER	P.P.M.	P.P.M. (TOT.)
1861-1900	47	9.4	1.88	4	0.8	0.16
1901-1922	34	6.8	1.36	1	0.2	0.04
1923-1945	36	7.2	1.44	6	1.2	0.24
1946-1967	20	4	0.8	14	2.8	0.56
1968-2001	18	3.6	0.72	8	1.6	0.32
TOT.	155	/	6.2	33	/	1.32

Table 8: CORIS - Frequency rates of *ringraziare di* and *ringraziare per*

CORIS	RINGRAZIARE DI	P.P.M.	RINGRAZIARE PER	P.P.M.
1980-2011	210	1.61	468	3.6

The findings show that the use of *grazie per* and *ringraziare per* has considerably increased in the most recent history of Italian while, at the same time, the use of traditional forms *grazie di* and *ringraziare di* has decreased. This is evidenced by comparing the total number of occurrences in p.p.m. from 1861 to 2001 for *grazie per* (0.24) and *grazie di* (1.96) with the number of occurrences of the two forms from 1980 to 2011, respectively 4.72 vs 2.23.

The situation is similar for *ringraziare di* vs *ringraziare per* where constructions with *di* have been found to have a considerably higher frequency from 1861 to 1945 than constructions followed by *per* (4.68 vs 0.44). Again, the gap gets narrower and narrower with the passing of time until the situation is turned upside down in more recent years; from 1980 to 2011 the relation between *ringraziare di* and *ringraziare per* is 1.61 vs 3.6. The significance tests' result show that the difference in the frequencies is indeed significant over time with chi square value for *grazie di* and *grazie per* in the DiaCORIS being 14.49, df value 4 and *p-value* 0.005893517. Notice that in the MIDIA Corpus the difference in the frequencies of *grazie di* and *grazie per* was on the contrary not significant, showing that in more recent times the gap in the distribution between the two forms is even wider than in earlier stages of Italian. The results are also significant for *ringraziare di* and *ringraziare per* where the chi square value is 25.10, df value is 4 and *p-value* is 0.00004793403.

The chi square test results have then proved that the difference in the distribution of the two forms is significant in the DiaCORIS with collocations with *di* being consistently more frequent. The log likelihood test (LL) will now give us the possibility to find out if the difference in the distribution of forms with *per* in the DiaCORIS and in the CORIS is significant. The results are the following: the LL value for total occurrences of *grazie per* constructions in the DiaCORIS and in the CORIS is 127.73 while for *ringraziare per* is 41.75. These results are extremely significant and they conclusively prove that the change in the increased use of constructions with *per* is as recent as 40-50 years.

In the examples below, excerpts [9], [10], [11], and [12] allow us to verify if the distribution of the two concurrent variables can be correlated with the context of use or other variables.

[9] Addio, Pinocchio, - rispose il cane; - mille **grazie di** avermi liberato dalla morte (Narrativa - 1883)

Farewell Pinocchio – the dog said – and thank you for saving me from death

[10] Domattina devo levarmi presto. E **grazie per** la bella compagnia (Narrativa – 1958)

I have to get up early tomorrow. Thank you for the good company

[11] Cari amici, **grazie di** essere qui [EPHEMOpuscoli 1980-2011)

Dear friends, thanks for being here

[12] **Grazie per** aver cercato in tutti i modi di ostacolarmi, capo! (MON2001_04 1980-2011)

Thanks for trying so hard to hinder me, boss!

The examples show once more that there does not seem to be any correlation between the choice of the preposition and the context of use which indicates that the change may be due to other sociolinguistic variables, such as an influence from the English *thanks/to thank for*. In the next section, I will analyse the distribution of *grazie/ringraziare di* and *grazie/ringraziare per* in the corpus of Italian films dubbed from (American) English which collects the dialogues of fifteen films distributed in Italy from 1964 to 2007.

2.4 Distribution in Italian dubbed films

I will now analyse the distribution of the two variants in a small corpus of Italian films dubbed from (American) English to assess which form collocates more strongly. The results of the investigations will be particularly relevant for discussions concerning the alleged influence from English via dubbing translations as causing the increase in the use of forms followed by *per* in real use Italian. Tables 9 and 10 show the raw frequency rates and the correspondent relative proportions (per thousand words) for occurrences of both constructions.

Table 10: Film Corpus – Frequencies of *ringraziare di* and *ringraziare per*

FILM CORPUS	RINGRAZIARE DI	RELATIVE	RINGRAZIARE PER	RELATIVE
1964-2007	0	0	5	0.03

Table 11: Film Corpus – Frequencies of *grazie di* and *grazie per*

FILM CORPUS	GRAZIE DI	RELATIVE	GRAZIE PER	RELATIVE
1964-2007	0	0	9	0.05

In the film corpus, there are no occurrences for constructions with *di*; the fact that occurrences with *per* have on the contrary been found shows that the communicative situation in question is represented thus excluding the possibility that a lack of occurrences may be due to a lack of representation in the corpus of the speech act of thanking somebody for something. More importantly, it evidences that, even in a relatively small corpus, variants with *per* are strongly preferred. Here below, [12], [13], and [14] are examples taken from the corpus (bold mine).

[12] Jack, voglio **ringraziarla per** quello che ha fatto (1994-1997)

Jack, I want to thank you for what you did

[13] **Grazie per** la collaborazione (1971-1975)

Thanks for your cooperation

[14] Signore, **grazie per** la vostra piacevole compagnia (1994-1997)

Sir, thanks for the lovely company

From the above excerpts, it can be seen how traditional constructions with *di* would have been equally acceptable. These results are particularly significant in relation to the hypothesis that the predominant use of constructions with *per* in Italian dubbed films may have played a fundamental role in boosting the increase of such forms in real use Italian. Such discussions are also supported by the DiaCORIS and CORIS results which conclusively established that the increase of *grazie/ringraziare per* constructions can be observed starting from the mid/late 20th century, which coincides with the coming of sound cinema and dubbing in Italy.

In the next section, I shall analyse the distribution of the two forms in three corpora of spoken Italian which collect samples of oral dialogues from 1965 to 2003.

2.5 Distribution in spoken Italian

I will now move on to investigate the distribution of *grazie/ringraziare di* and *grazie/ringraziare per* in three corpora of spoken Italian which collectively amount for about 1 million words. Because these three corpora are different in size, the log-likelihood test (LL) will be performed as the significance test alongside the chi square test within each corpus for both constructions. Table 12 shows the results for *grazie di* and *grazie per* while table 13 shows the results for *ringraziare di* and *ringraziare per*.

Table 12: Frequency rates of *grazie di* and *grazie per* in spoken Italian from 1965 to 2003

	GRAZIE DI	P.P.M.	GRAZIE PER	P.P.M.
STAMM. (1965)	1	9.99	0	0
LIP (1990-1992)	4	8.18	3	6.13
C-ORAL-ROM (2000-2003)	1	3.21	24	77.03

Table 13: Frequency rates of *ringraziare di* and *ringraziare per* in spoken Italian from 1965 to 2003

	RINGRAZIARE DI	P.P.M.	RINGRAZIARE PER	P.P.M.
STAMM. (1965)	0	0	0	0

LIP (1990-1992)	7	14.31	18	36.80
C-ORAL-ROM (2000-2003)	6	19.26	3	9.63

The results show that no occurrences for *grazie per* have been found in the older corpus of spoken Italian (Stammerjohann 1965) while the form registered a high frequency rate in the most recent corpus (2000-2003). The LL value for these two results is 13.38 which is extremely significant as the probability of this result happening by chance is less than 1%. Similarly, in the Stammerjohann (1965) no occurrences for *ringraziare per* have been found and the LL value between the Stammerjohann and the LIP is 6.70, which, again, is extremely significant (99% certainty that the difference between the two results is not due to chance). At the same time, the chi square value for occurrences of *grazie/ringraziare di* and *grazie/ringraziare per* within the LIP is 0.97 and the *p*-value is 0.3247558, which is not significant, while the chi square value is 12.29 with a *p*-value of 0.0004545219 in the C-ORAL-ROM. This means that the difference in the distribution of the two forms in spoken Italian is extremely significant in the most recent corpus thus proving that in contemporary spoken Italian forms with *per* are overall preferred to forms with *di*.

3. Conclusions

This article investigated the diachronic trend of two Italian constructions, *grazie/ringraziare di* ‘thanks/to thank of’ (thanks/to thank for) and *grazie/ringraziare per* ‘thanks/to thank for’. The main research hypothesis is that, although both forms are equally acceptable and constructions with *di* have been historically more established than forms with *per*, in recent times *grazie/ringraziare per* have increasingly been used in Italian. A number of authors (e.g., Renzi 2000, Alfieri et al 2008) have also hypothesised that such an alleged increase may be due to an influence from the English *thanks/to thank for*; however, both the claims of an increase in real use Italian of *grazie/ringraziare per* and the alleged influence from English had not been substantiated by any empirical investigation. On the contrary, this study used empirical data such as language corpora of written, spoken and dubbed Italian, to investigate the diachronic distribution of *grazie/ringraziare di* and *grazie/ringraziare per* in three language varieties (written, spoken, dubbed) so as to obtain a full and detailed account of the trend of these constructions through the history of Italian.

The etymological and lexicographic analyses conducted in § 2.1 have reported occurrences for *ringraziare di* while no occurrences for constructions with *per* have been found thus suggesting that the preposition *di* historically collocated more strongly than *per*. Diachronic quantitative investigations have been carried out across corpora of authentic written and spoken Italian from 1200 to 2011 to empirically support this finding and to assess whether and when the users’ choices

have shifted towards forms with *per*. The results confirmed that, while constructions with *per* were extremely rare in older stages of Italian, the use of *grazie/ringraziare per* has remarkably increased over time, to the point that the frequency of occurrence of such constructions is higher in contemporary Italian than traditional forms with *di*. Moreover, the diachronic analysis of the distribution of both forms has shown that, while the difference in the frequency distributions between *ringraziare di* and *ringraziare per* has constantly been significant (with forms with *per* being consistently less frequent), the difference in the frequency distributions of *grazie di* and *grazie per* has only been found significant from 1861 (with forms with *per* being less frequent). Such findings were particularly relevant when compared with data in the corpus of dubbed Italian and after the coming of dubbing in Italy, where a clear dominance of forms with *per* has been found instead.

These investigations empirically proved that a change in the Italian language has taken place and indicate that, in absence of other sociolinguistic variables, such as the context of use or the expressed function, strong positive correlations exist with an influence from English via dubbing. Specifically, it was shown here that: 1) forms followed by *di* were historically preferred, 2) in dubbed Italian, forms with *per* are strongly preferred and 3) the use of constructions with *per* have substantially increased after the introduction of dubbing in Italy to the point that in contemporary real use Italian they are in fact more frequent than constructions with *di*.

The study did not try to claim that dubbing is the *only* direct cause for the increase in the use of *ringraziare per* and *grazie per* as factors such as globalisation, the Internet, and the predominance of English as lingua franca are acknowledged as major sources of influence. However, if on the one hand the role of dubbing cannot be isolated, on the other, positive correlations have been found, therefore suggesting that the role played by dubbing cannot be excluded. In other words, the results of this study show that such a concurrent role, though may not be measurable, is however demonstrable.

¹ The films are Mary Poppins 1964, Dr Strangelove 1964, A Space Odyssey 1968, The Andromeda Strain 1971, Young Frankenstein 1974, Monty Python and the Holy Grail 1975, Shining 1980, Back to the Future 1985, Life of Brian 1991, Pulp Fiction 1994, Apollo 13 1995, Titanic 1997, the Big Kahuna 1999, Donnie Darko 2001, 300 2007

References

Alfieri, Gabriella, Motta Daria & Rapisarda Maria. 2008. La fiction. *Gli italiani del piccolo schermo. Lingua e stili comunicativi nei generi televisivi* ed. by Gabriella Alfieri & Ilaria Bonomi, 235-340. Firenze: Franco Cesati Editore.

- Babbie, Earl R. 2013. *The Practice of Social Research*. 13th ed. Belmont, CA: Cengage Learning.
- Cinetel. 2016. *Dati annuali cinema*. Retrieved from <http://www.anica.it/web/ricerche-e-studi/dati-annuali-cinema>. Last accessed on 13/06/2016.
- Eckert, Penelope. 2000. *Linguistic variation as social practice*. Oxford: Blackwell.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12: 453-76.
- Giles, Howard. 1984. The dynamics of speech accommodation. *International Journal of the Sociology of Language*, Special issue, 46: 1-155.
- Giles, Howard and Powesland, Peter F. 1975. *Speech Style and Social Evaluation*. London: Academic Press.
- Giles, Howard, Coupland, Justine & Coupland, Nikolas, eds. 1991. Accommodation Theory: Communication, Context, and Consequence. *Contexts of Accommodation*, 1-68. New York, NY: Cambridge University Press.
- Labov, William. 1994. *Principles of linguistic change, vol. 1: Internal factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistic change, vol. 2: Social factors*. Oxford: Blackwell.
- Labov, William. 2010. *Principles of linguistic change, vol. 3: Cognitive and cultural factors*. Oxford: Wiley-Blackwell.
- McEnery, Tony and Wilson, Andrew. 2001. *Corpus Linguistics. An Introduction*. Edinburgh: Edinburgh University Press.
- Milroy, James. 1992. *Linguistic variation and change*. Oxford: Blackwell.
- Milroy, James. 2002. Introduction: Mobility, contact and language change -working with contemporary speech communities. *Journal of Sociolinguistics* 6: 3-15.
- Renzi, Lorenzo. 2000. Le tendenze dell'italiano contemporaneo. Note sul cambiamento linguistico nel breve periodo. *Studi di lessicografia italiana*, 17: 279-319.
- Renzi, Lorenzo, Salvi, Giampaolo, and Cardinaletti, Anna, eds. 1991. *Grande grammatica italiana di consultazione vol. II*. Bologna: Il Mulino.

Serianni, Luca. 2010. Sulla reggenza di *grazie*. Accademia della Crusca. Retrieved from <http://www.accademiadellacrusca.it/it/lingua-italiana/consulenza-linguistica/domande-risposte/reggenza-grazie>. Last accessed on 01/07/2016.

Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.

Corpora

C-ORAL-ROM (Integrated Reference Corpora for Spoken Romance Languages). 2005. Cresti, Emanuela and Moneglia, Massimo, eds. Amsterdam/Philadelphia: John Benjamins.

CORIS (Corpus di italiano scritto). Rossini Favretti, Raffaella, Tamburini, Fabio & De Santis, Cristiana. 2002. A corpus of written Italian: a defined and a dynamic model. *A Rainbow of Corpora: Corpus Linguistics and the Languages of the World* ed. by Andrew Wilson, Paul Rayson & Tony McEnery, 27-38. Lincom-Europa: Munich. Retrieved from <http://corpora.dslo.unibo.it/TCORIS/> on 29/09/2015.

DiaCORIS (Corpus diacronico di italiano scritto). Onelli, C., Proietti, Domenico, *et al.* 2006. The DiaCORIS project: a diachronic corpus of written Italian. *Proceedings of 5th International Conference on Language Resources and Evaluation (Genova, 22-28/5/2006)*, Genova: LREC, 1212-1215. Retrieved from http://corpora.dslo.unibo.it/DiaCORIS/_on29/05/2015.

LIP (Lessico di Frequenza dell'Italiano Parlato). De Mauro, Tullio, Mancini, Federico, Vedovelli, Massimo, & Voghera, Miriam, eds. 1993. Milano: Etaslibri. Retrieved from <http://www.parlaritaliano.it/index.php/it/volip> on 29/05/2015.

MIDIA (Morfologia dell'italiano in Diacronia). Gaeta Livio, Iacobini Claudio, Ricca, Davide, Angster, Marco, De Rosa, Aurelio & Schirato Giovanna, "MIDIA: a balanced diachronic corpus of Italian", comunicazione presentata alla 21st International Conference on Historical Linguistics (Oslo, 5-9 agosto 2013). Retrieved from www.corpusmidia.unito.it on 10/11/2016.

OVI Corpus (Opera del Vocabolario Italiano). Istituto del Consiglio Nazionale delle Ricerche. Retrieved from www.gattoweb.ovi.cnr.it last accessed on 10/11/2016.

Stammerjohann. Tucci, Ida and Signorini, Sabrina. 2004. Il restauro e l'archiviazione elettronica del primo corpus di italiano parlato: il corpus Stammerjohann. *Atti delle Giornate del Gruppo di Fonetica Sperimentale - XIV*, Viterbo, 4-6 dicembre 2003. Roma: Esagrafica, 119-126. Retrieved from <http://lablita.dit.unifi.it/corpora/imdi/stam/> 29/09/2015.

Dictionaries

- Battaglia, Salvatore. 1961-2002. *Grande dizionario della lingua italiana*. Torino: UTET = GDLI.
- Cannella, Mario, ed. 2007. *Lo Zingarelli 2008. Vocabolario della lingua italiana di Nicola Zingarelli con CD-ROM*. Bologna: Zanichelli = Zing. 2008.
- Cortelazzo, Mario and Zolli, Paolo, eds. 2008. *Il nuovo etimologico. Dizionario etimologico della lingua italiana. Con CD-ROM*. Bologna: Zanichelli = DELI.
- De Mauro, Tullio. 2000. *Dizionario della Lingua Italiana De Mauro. Versione elettronica*. Torino: Paravia.
- Devoto, Giacomo and Oli, Giancarlo. 2013. *Il Devoto-Oli. Vocabolario della lingua italiana 2014. Versione elettronica* ed. by Luca Serianni and Pietro Trifone. Firenze: Le Monnier = Devoto-Oli 2014.
- Grande dizionario di Italiano*. 2013. Milano: Garzanti Linguistica = GDI. Retrieved from www.garzantilinguistica.it 29/05/2015.
- Migliorini, Bruno. 1950. Appendice al Dizionario moderno. *Dizionario moderno* (nona edizione) ed. by Panzini, Alfredo, 765-997. Milano: Hoepli.
- Nocentini, Alberto, ed. 2010. L'Etimologico. *Vocabolario della lingua italiana. Versione elettronica*. Firenze: Le Monnier.
- Panzini, Alfredo. 1905. *Dizionario Moderno delle parole che non si trovano nei dizionari comuni*. Milano: Hoepli.
- Tommaseo, Niccolò and Bellini, Bernardo. *Dizionario della lingua italiana con oltre centomila giunte a precedenti dizionari. Raccolte da Niccolò Tommaseo, Giuseppe Campi, Giuseppe Meini, Pietro Fanfani e da molti altri distinti filologi e scienziati. 1861-1879. Lo Zingarelli 2008. Vocabolario della lingua italiana di Nicola Zingarelli con CD-ROM*. 2007 ed. by Mario Cannella. Bologna: Zanichelli.
- Vocabolario Degli Accademici Della Crusca*. Quarta edizione 1729-1738. Firenze: Domenico Maria Manni.
- Vocabolario Degli Accademici Della Crusca*. Quinta edizione 1863-1923. Firenze: Tipografia Galileiana.
- Vocabolario Treccani*. Roma: Istituto della Enciclopedia Italiana. Retrieved from www.treccani.it/vocabolario/ 29/05/2015.