Greek diglossia in the 20th century: A historical corpus linguistics approach

Georgia Fragaki & Dionysis Goutsos
(University of Athens)

Structure of the paper

- The questions of Greek diglossia

- Historical corpus linguistics and recent language change

- *Diachronic Corpus of Greek of the 20th century*

- Data and method

- Findings: Grammatical items in diachrony

- Findings: Text types

- Conclusions

The questions of Greek diglossia I

"what is societally recognized as diglossia is not directly related to how actual people implement the relevant contrasts" (Daltas 1993: 348)

> *what actual people did in specific contexts*

(cf. Alexiou 1982: 156: "what is the actual spoken and written usage of Greek today", Mackridge 2009: 29: "actual language used in Greece")

The questions of Greek diglossia II

A series of unanswered questions:

- how many language varieties?

      Fergusonian diglossia: K (High) vs. D (Low)

      Mirambel's (1937) five-part distinction for the first half of 20th c.:
- katharévousa
- 'mixed' (scientific and political texts, news reports)
- 'colloquial' (everyday urban, journalism, non-fiction)
- demotic (everyday rural, dialectal variation)
- 'hairy' (standardization attempted by demoticists)

The questions of Greek diglossia II

- how many language varieties? (cont.)

Alexiou (1982: 162-163): "D and K [...] cannot be divided into two absolutely watertight compartments, each with a separate set of syntactic, morphological and lexical structures [...] [Mirambel's] categorisation suggests a fixed and static situation which is misleading [...] shades and fluctuations between katharevousa and demotic"

Kazazis (2001: 293): "the binary opposition between High and Low is merely a convenient fiction [...] the span between these putative opposites constitutes a continuum"

Holton (2002: 171): "many forms of K and many forms of D"

The questions of Greek diglossia III

- what was the range of registers?

Mackridge (2009: 29): "actual language used in Greece covered a continuum of linguistic registers ranging from 'pure' demotic to 'extreme' katharevousa, with hybrid varieties in between".

Holton (2002: 171): "K was the language of officialdom, the law, science and much of the educational system; it also had a considerable influence on journalism and the broadcast media. D was the language of oral communication, informal writing, poetry and novels".

The questions of Greek diglossia IV

- is it a question of spoken vs. written?

> Horrocks (2010: 459): "the written styles of the early 20th century were just too remote from ordinary speech to be learned effectively by the majority".
>
> Mackridge (2004): "a fundamental misunderstanding of Greek diglossia, which has usually been presented as a conflict between spoken and written usage"
>
> Petrounias (1978: 201): "common dialect evolved in the 19th c. which was used both in oral and written communication" (Pre-Common Modern Greek)

The questions of Greek diglossia V

- when did changes happen?

> Mackridge (2010: 566): "By 1900 the use of Ancient Greek had completely died out except as a scholarly exercise".

> Mackridge (2009: 313): "In poetry and fiction written from the 1950s onwards [...] the literary language became more standardized [...] The language question had long been solved in literature, which waited for other discourses to catch up".

The questions of Greek diglossia VI

- how did the public influence the private?

Hudson (2002): sociological (contextual) vs. grammatical (linguistic) aspects of diglossia

Daltas (1993: 341-342): "Ferguson's paradigm should be taken to be about the <u>public</u> aspect of diglossia as opposed to the <u>private</u>, about the legal system and not about the way people respond to it in actual life, though <u>the former influences the latter</u> in complex, dynamic ways".

The questions of Greek diglossia VII

Lack of evidence from use:

Goutsos et al. (1994: 216): "arguments on linguistic issues are still based on an extremely small amount of data"

Iordanidou (2002a, b): "lack of backup from corpora"

> development of a diachronic corpus of Greek

Historical corpus linguistics and recent language change I

- Development of diachronic corpora e.g. *Helsinki Corpus* (1,5 m., 730-1710), *COHA* (400 m., 1810-2009)
- Corpus "families" e.g. *Lancaster-Oslo-Bergen Corpus* (including BLOB-1931: 1928-1934, LOB:1961 and F-LOB: 1991)
- Synchronic, monitor corpora e.g. *BNC* (1960-now)

Phenomena studied:

- vocabulary changes (Baker 2011)
- grammatical change (Leech et al. 2009)
- morphology (Baayen & Renouf 1996, Fischer 1998, Duguid 2010)
- phraseology (Davies 2012)
- cultural changes (Baker 2010, Marchi 2010, Partington 2012).

Historical corpus linguistics and recent language change II

Historical corpus linguistics: a convergence of disciplines:

- historical sociolinguistics (Romaine 1982, Hernandez-Campoy & Conde-Silvestre 2012, Nevalainen & Raumolin-Brunberg 2003)
- historical pragmatics (Jucker 1995, Jucker & Taavitsainen 2010)
- grammaticalisation (Hopper & Traugott 1993/2003)
- corpus linguistics on recent language change (Mair 2009)

- A variationist view of language change (Milroy 1992: 123)
- Attention "to the role played by textual and discourse factors across the centuries" (Rissanen et al. 1997a: v, 1997b: v)

Diachronic Corpus of Greek of the 20th century

Research project "Aristeia I" (Excellence)

Aims:

a) to examine the issues involved in the compilation of a diachronic corpus of Greek, including the availability of data/text types and representativeness

b) to collect data for a diachronic corpus of Greek of the 20th century; to complement the existing *Corpus of Greek Texts* (1990-2010)

c) to analyze the corpus with a view to drawing basic conclusions on linguistic change

Data and method I

Selection of data from *GreekCorpus20*:

Genres	Text types	Number of words
Spoken news	Newsreels	78,441
Public speeches	Parliament, Academic, Other	339,194
Conversation	Film scripts	208,206
Literature	Novels, Short stories, Poetry, Drama	1,086,687
Academic	Humanities, Social/Finance, Science	1,044,200
Popularized	Non-fiction	285,252
Private	Letters	178,720
TOTAL		**3,220,700**

Data and method II

Raw and normalised (per 10.000 words) frequencies of seven

grammatical items: *ἵνα* and the pairs: *ποῖος* vs. *ποιός/ποιος,*

διά vs. γιά/για, εἰς vs. *σέ/σε,* (incl. variants e.g. δι', γι', σ', στ*)

- seven text types: emphasis on *speech-like* (private letters),

speech-based (public speeches) and *speech-purposed* (films,

drama, newsreels) text types (cf. Culpeper & Kytö 2010)

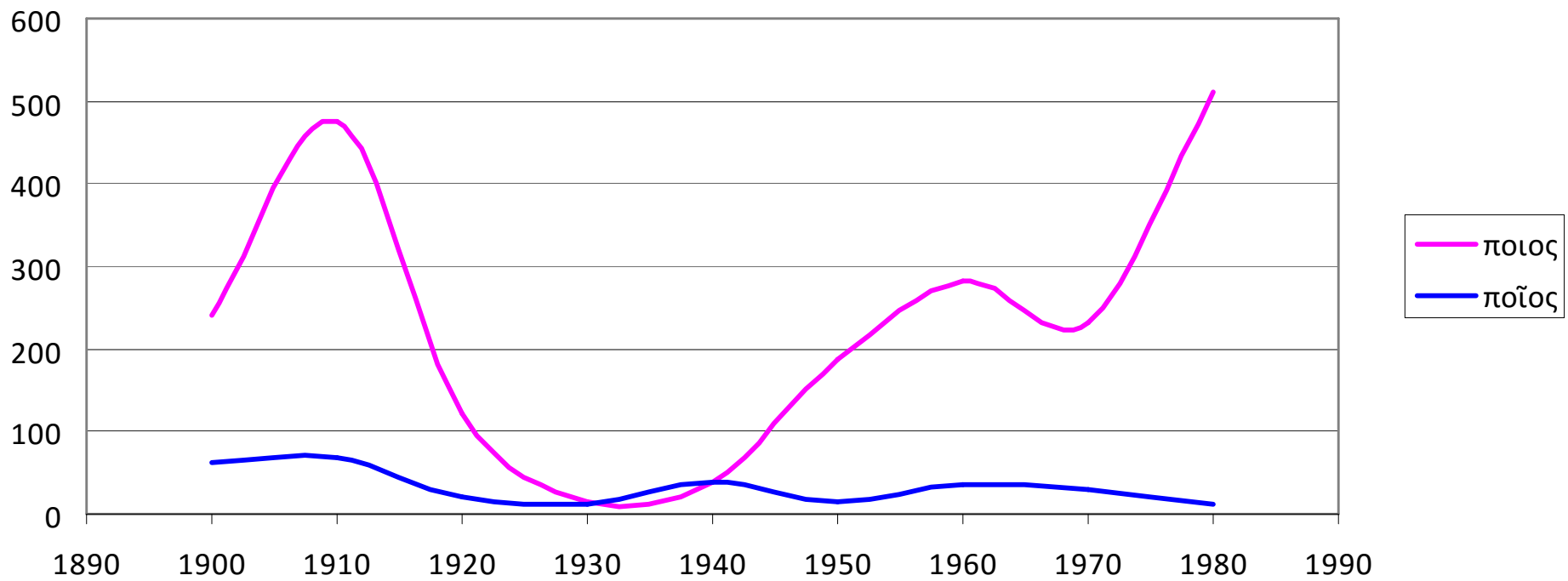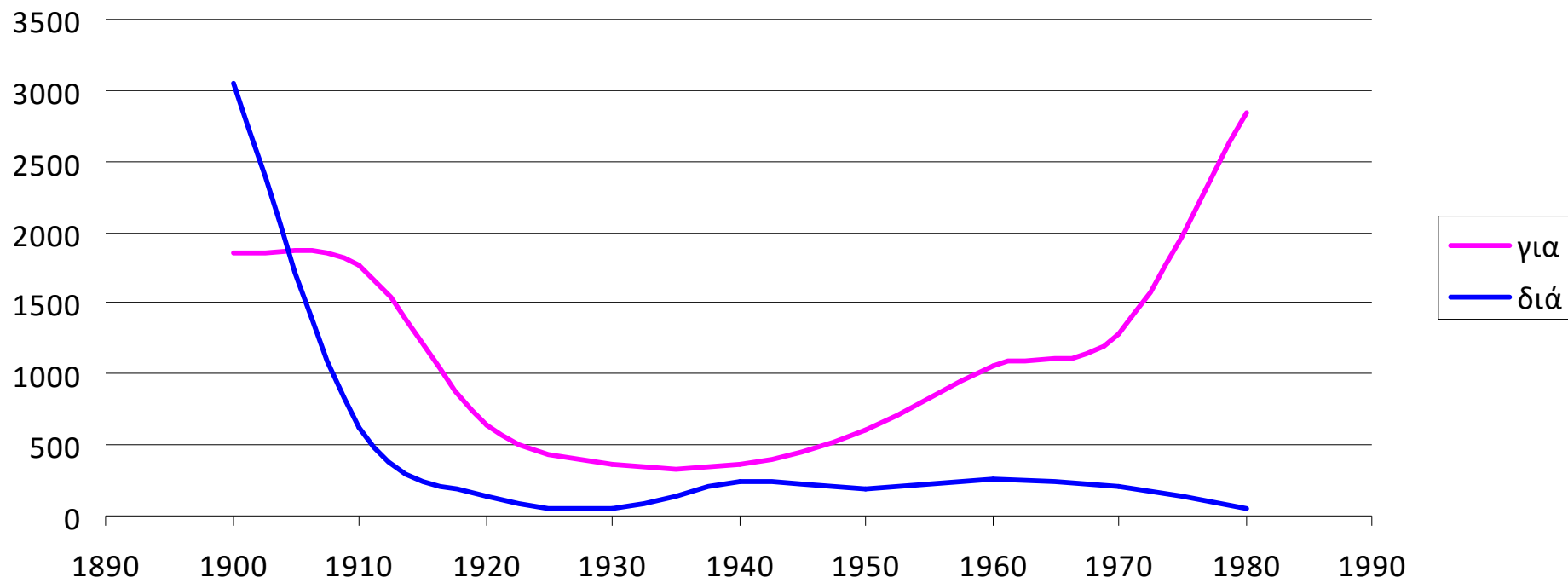- data from nine decades (1900-1989)

Grammatical items in diachrony I

Grammatical items in diachrony II

ἵνα
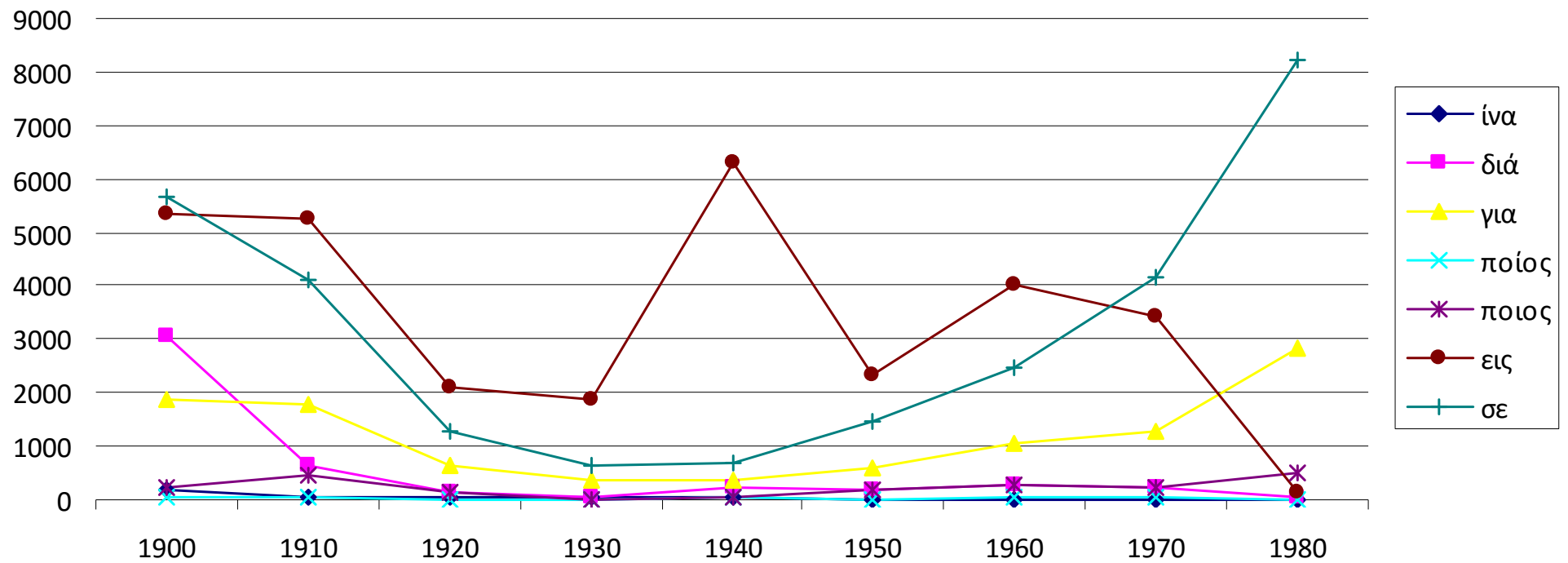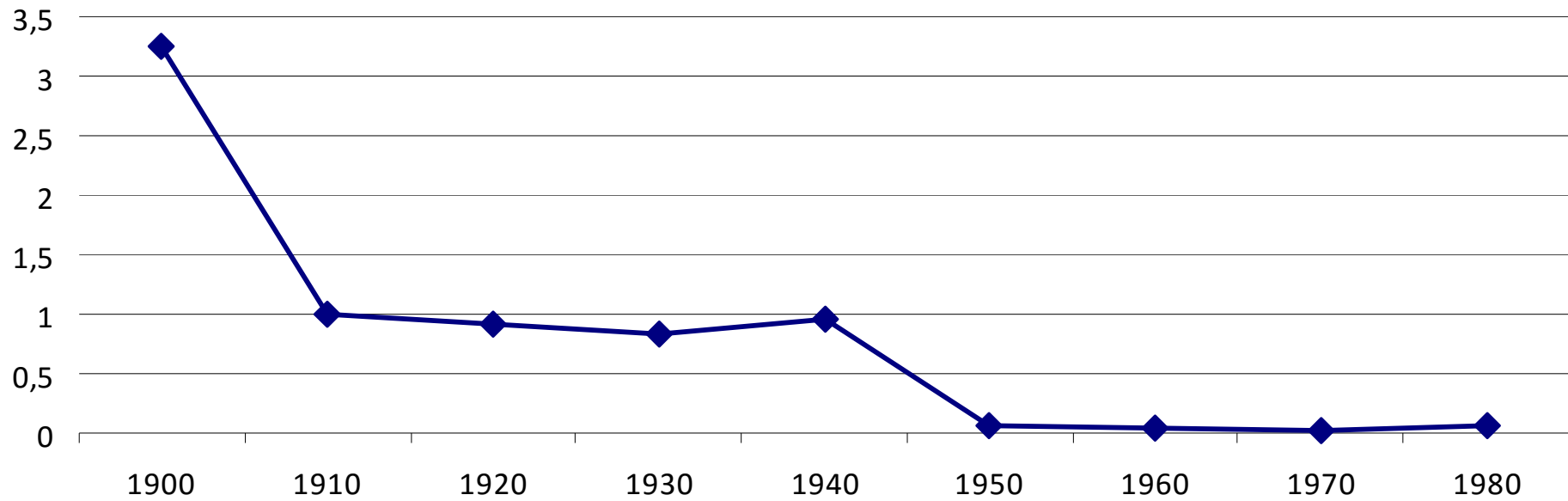
Grammatical items in diachrony III

Grammatical items in diachrony IV

Grammatical items in diachrony V
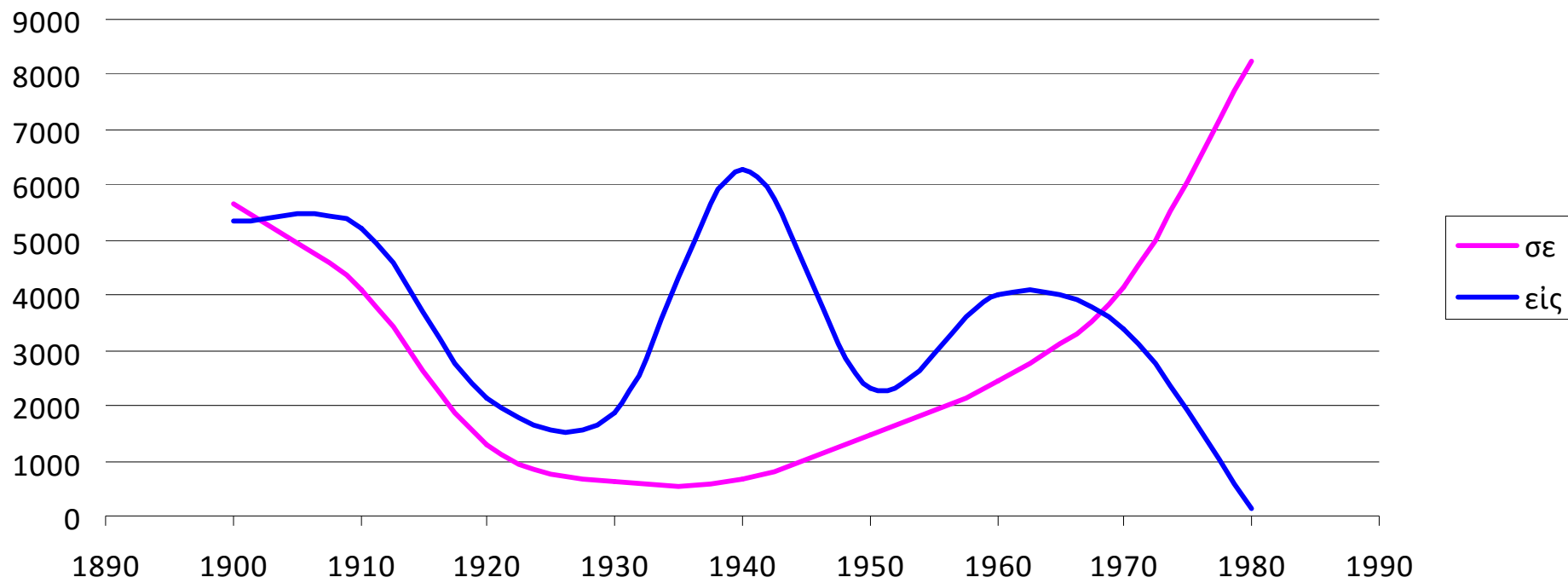
Text types I

Newsreels

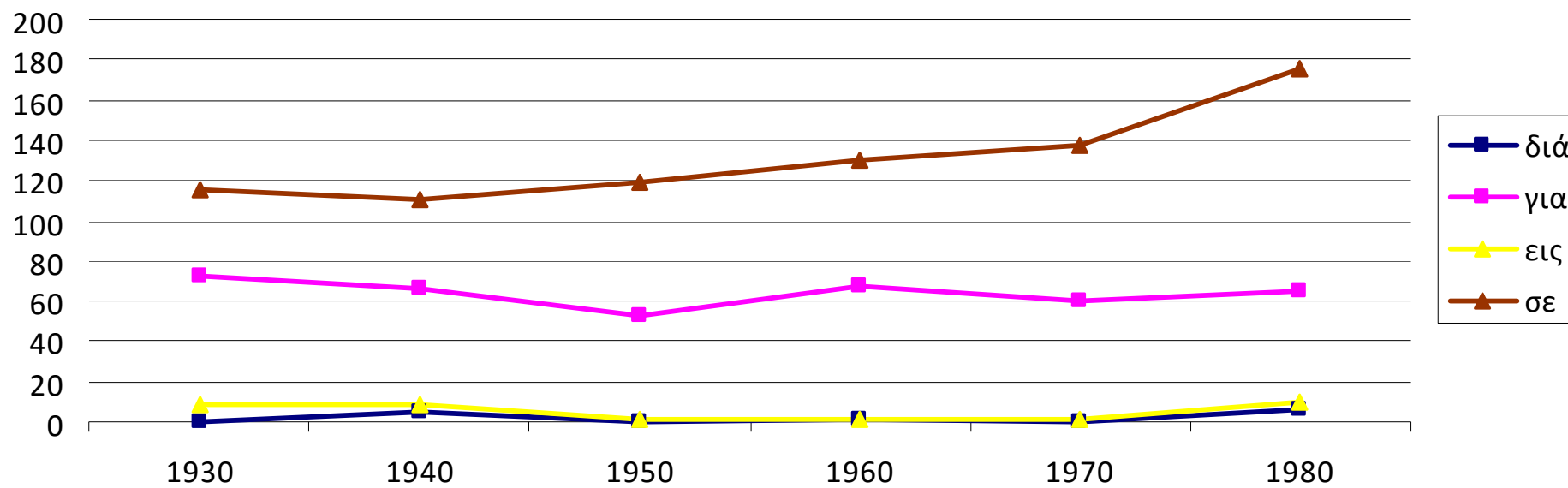| Legend |
|---|
| διά |
| για |
| εις |
| σε |
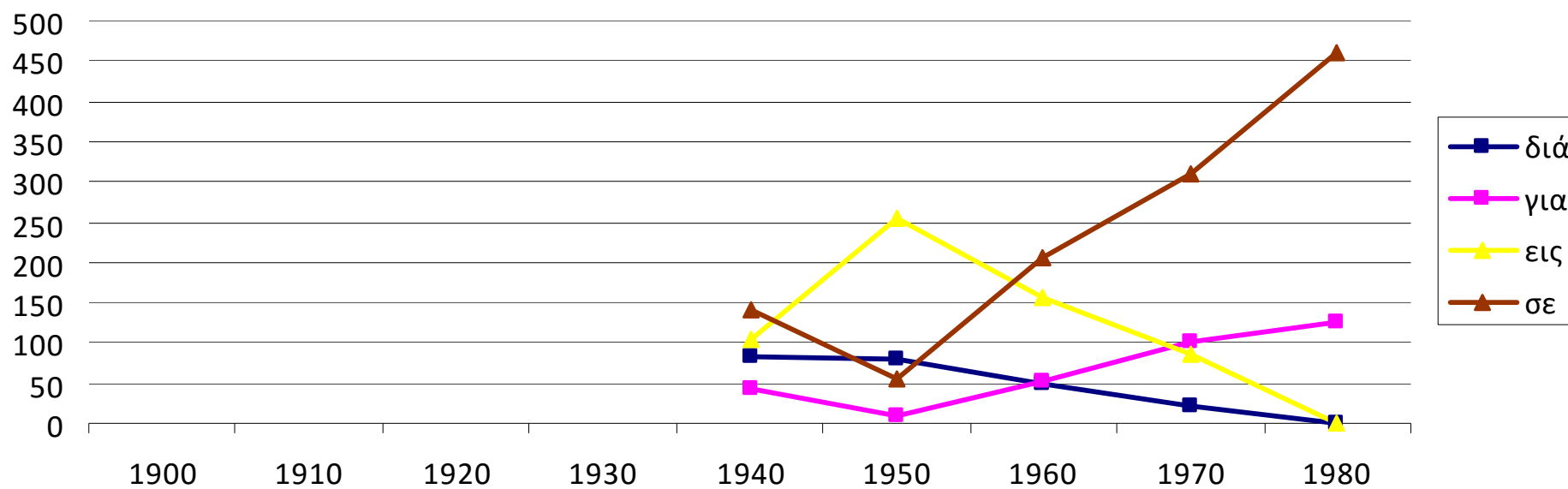
Text types II

Speeches

Legend:
- διά
- για
- εις
- σε
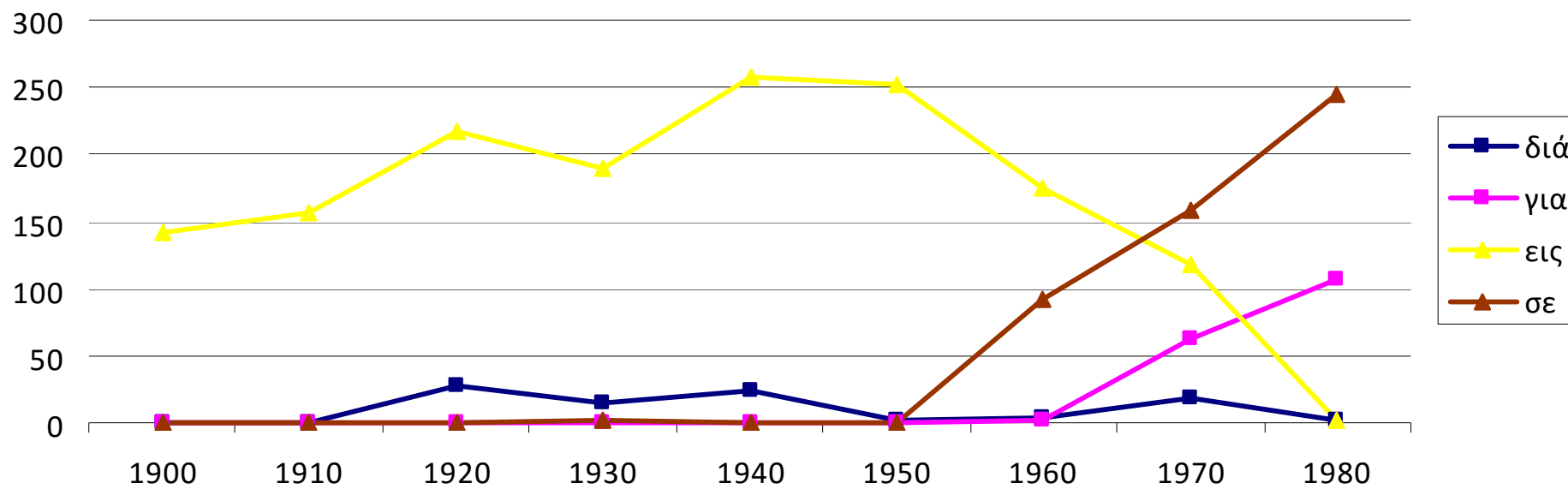
Text types III

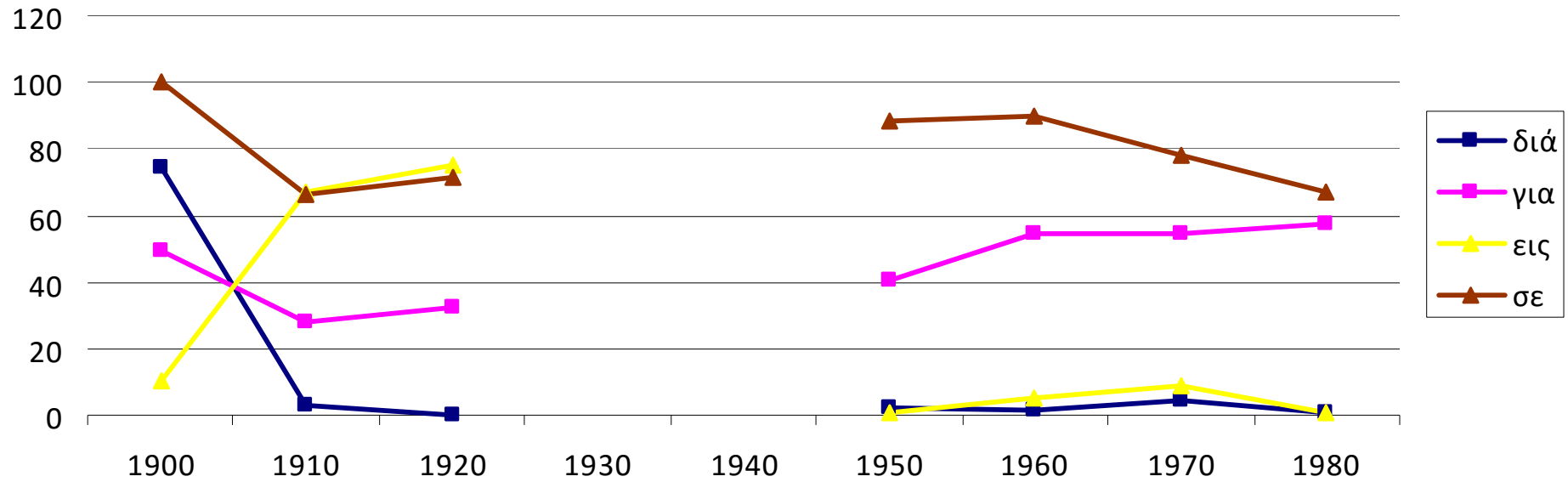Films

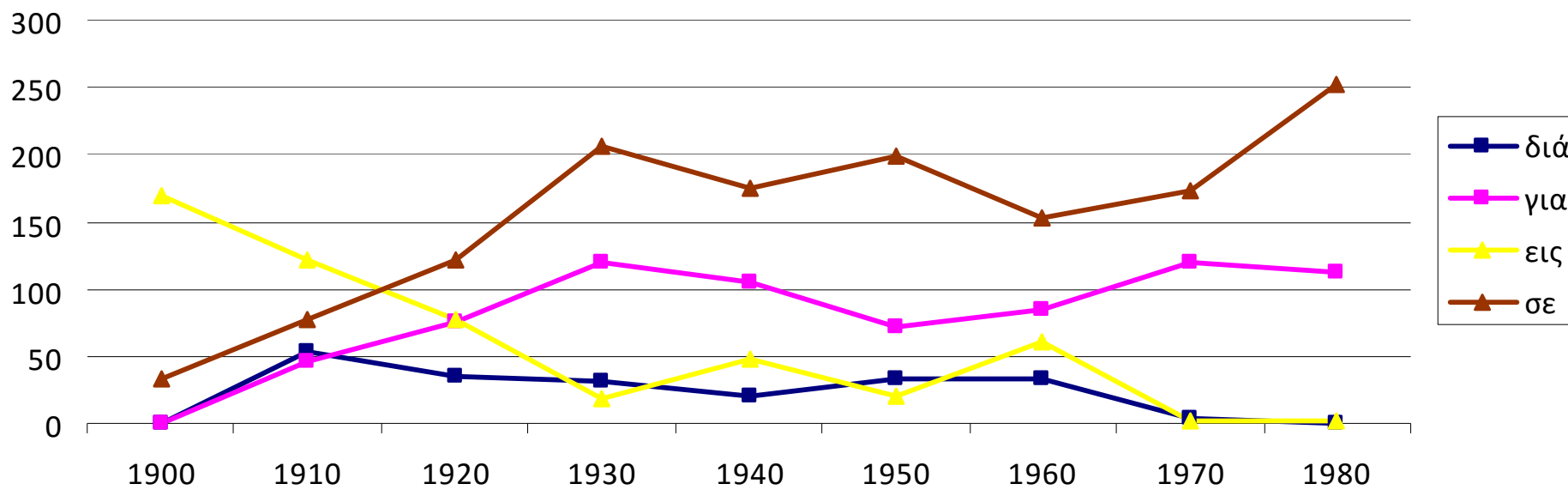# Text types IV

Academic

Legend: διά, για, εις, σε
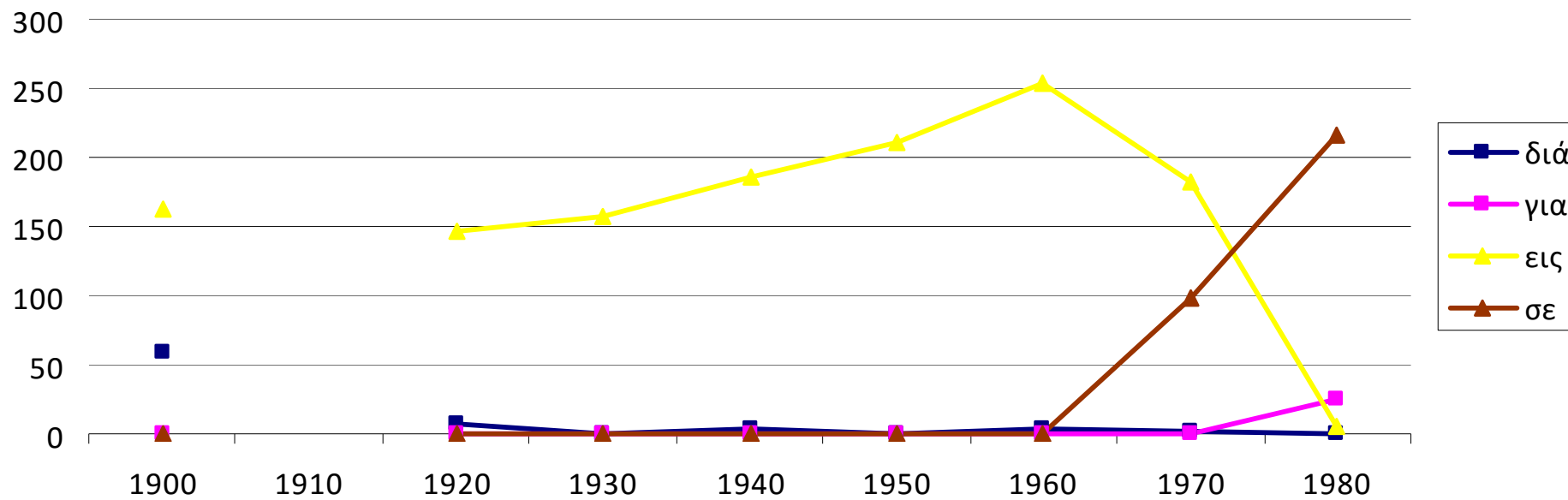
Text types V

Literature

Text types VI

Private letters

Conclusions I

a variationist view of language change on the basis of the thoroughly attested role of frequency (Schneider 2004):
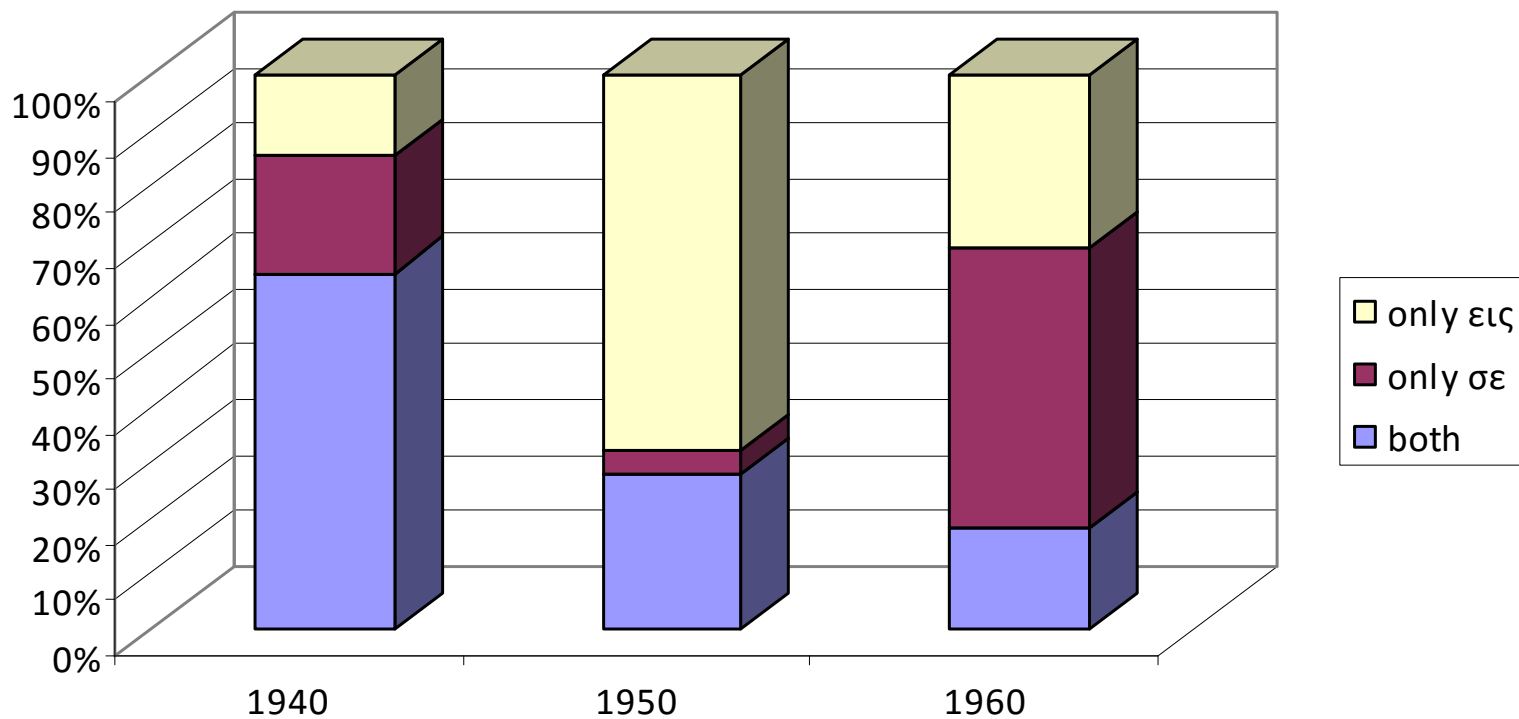
U-curve of D variants in Greek diglossia

"Roller-coaster" pattern of K variants > stereotypical (in Labov's sense), emblematic use

Formal vs. informal text types: conservative with sudden changes vs. gradual, "normal" language change

Conclusions II

Co-existing variants within text types/individual texts?

- ες vs. σε in newsreels 1940s-1960s

Conclusions III

Co-existing variants within text types/individual texts?

- εις vs. σε in newsreels 1940s-1960s

Αι μηχανοκίνηται δυνάμεις του ελληνικού στρατού παρελαύνουν <u>εις</u> επιβλητικούς σχηματισμούς και εμπνέουν <u>εις</u> τας μυριάδας του λαού που τις παρακολουθεί το αίσθημα της ασφαλείας και της υπερηφανείας. […] Χιλιάδες των θεατών χειροκροτούν και επευφημούν τους άνδρας αυτούς, οι οποίοι, με τους υπερόχους αγώνας των, έδωσαν <u>στην</u> Ελλάδα μας την μεγάλην και ωραίαν της νίκην (SRF01-1940-0002)

Μετά την άφιξιν <u>εις</u> την πόλιν, τα παιδιά δέχονται την ευλογίαν της Εκκλησίας. Προσφέρονται <u>εις</u> τα παιδιά διάφορα δώρα. Τα παιδάκια φθάνουν <u>εις</u> παιδόπολιν της Θεσσαλονίκης. Με τη χαρά ζωγραφισμένη <u>στα</u> πρόσωπά των παίζουν με τα δώρα που τους προσεφέρθησαν […]. (SRF01-1950-0019)

<u>Στον</u> αερολιμένα του Ελληνικού συνεχίζεται με εντατικότατο ρυθμό <u>στα</u> πλαίσια του Τεχνικού Προγράμματος του 1969 η ανέγερσις νέων κτηριακών εγκαταστάσεων [….] <u>Στο</u> Τεχνικό Πρόγραμμα του 1969 προβλέπεται […] δια της χρησιμοποιήσεως ραντάρ <u>εις</u> τον έλεγχον της εναερίου κυκλοφορίας. […] να ανταποκρίνεται ανέτως <u>εις</u> τας υποχρεώσεις του ως διεθνούς αεροπορικού κόμβου και <u>εις</u> τας ανάγκας διακινήσεως εμπορευμάτων και επιβατών (SRF01-1960-0050)

Conclusions IV

Co-existing variants within text types/individual texts?

- εις vs. σε in private letters 1900s-1920s:

only 4% have both types > individual preference

- διά vs. για in literature 1900s-1910s:

most texts have both types > stylistic device:

> Μία εξ αυτών κατελήφθη υπό σπασμωδικού ακαταβλήτου γέλωτος και ενώ <u>δια</u> της μιας χειρός εκράτει τα στήθη της, <u>δια</u> της άλλης εκτύπα τον ώμον της Λενιάς […] συζητούντας και λύοντας τας διαφοράς των ενώπιον των προεστών, "<u>για</u> να παίρνη πράξι" […] Ήτο τότε πολύ μικρός· επειδή δε κάθε φορά εζήτει να τον σηκώνουν <u>για</u> να παίζη την καμπάνα, αυτός κατ' ουδένα λόγον να πλησιάση άλλους ανθρώπους. Ούτε <u>για</u> να μεταλάβη δεν έστεργε να καταιβή στο χωριό. […] Ο Μανώλης εγέλασεν, αλλ' εγέλασε μάλλον <u>δια</u> την μωρίαν εκείνου του νέου (WFB08-1910-0004)

Conclusions V

Indications about the questions of diglossia:

- spoken vs. written?

- when did changes happen?

- how did the public influence the private?

A full-scale investigation of variants is expected to contribute to

> an understanding of the range of registers and language varieties

> an informed view on standardisation

www.greekcorpus20.phil.uoa.gr