Some Reflections on Reference in the Teaching of French as a Foreign Language Context

Olympia Tsaknaki
Department of French, Aristotle University of Thessaloniki

Abstract
Grammatical and lexical cohesion are necessary conditions to guarantee a text’s coherence (Halliday and Hasan 1976). Grammatical cohesion is classified into four types of relation: reference, conjunction, substitution, and ellipsis. This study aims to investigate the use of the cohesive device of reference in written discourse produced by users/learners of French as a foreign language, to explore the kind of references they prefer and to highlight the difficulties they come across. Our study is based on a learner corpus composed of written productions of Greek-speaking users/learners of French who participated in National Foreign Language Exam System (Κρατικό Πιστοποιητικό Γλωσσομάθειας, ΚΠΓ). We used a sample of texts produced by candidates who participated in the examination willing to be certified at the levels B (Independent user) and C (Proficient user) of language proficiency according to the 6-level scale of the Common European Framework of Reference for Languages (Council of Europe 2001).

Keywords: Reference, anaphora, cataphora, National Foreign Language Examinations (ΚΠΓ), French as a foreign language

1. Introduction
The ultimate purpose of the building up of the communicative competence is the attainment of a successful communication in any social context. Written communication requires the elaboration of clear, cohesive, well-structured texts without ambiguities that may prevent understanding. Well-formed texts are coherent texts, in other terms, texts which “must be semantically well-constructed and coherent linguistic units” (Halliday & Hasan 1976).

In the context of searching for, and achieving, the coherence of a text, the rule of repetition plays a prominent role. Recursive elements should be included in the text as a mandatory condition attached to the homogeneous linear development. The repetition of these units permits to link clauses or sentences (Charolles 1978 : 14-15).
Reference is the set of grammatical resources that allow the speaker to indicate whether something is repeated from somewhere earlier in the text, or whether it has not yet been repeated in the text (Thompson 2004). Anaphoric and cataphoric referencing are related cohesive devices with a different direction. They are both subsumed under the generic term called *endophora* (Maingueneau 2000: 172). *Anaphora* is the term which describes the relation of elements that contribute to the coherence of discourse and organization of the text when these elements follow the linguistic referent in the discourse. On the other hand, in the case of *cataphora*, these elements precede the referent (Riegel, Pellat & Rioul 1994: 612).

There has been already a focus on anaphora by linguists working on applied or corpus linguistics because anaphoric phenomena can provide us information on syntactic description of languages and important clues to Natural Language Processing. They can tell us "how discourse is constructed and maintained – how linguistic patterning above and beyond the sentence is arranged" (Botley & McEnery 2000: 3).

The purpose of this study is to make a first attempt to explore whether, and to what extent, reference is used adequately in written discourse of users/learners of French as a foreign language. Their sensibilization in the use of the appropriate elements could become a part of the written text analysis in classroom in order to help them to cope with difficulties, improve their skills and achieve the best results.

2. The *Common European Framework of Reference for Languages* (2020)

The user/learner of a foreign language is nowadays considered as a social agent. Since the objective of the contemporary curricula is the development of users/learners’ communication skills, language awareness, intercultural awareness and the ability to interact and mediate, the development of linguistic, sociolinguistic and pragmatic competence needs to act complementarily in the classroom of the foreign language. Coherence and cohesion, together with flexibility to circumstances, turntaking and thematic development, have been mentioned as one of the aspects of discourse competence in the *Common European Framework of Reference for Languages* (Council of Europe 2001: 124-125) and they continue to be key concepts in the more recent Companion Volume (Council of Europe 2020: 140-141).

At the C2 level, users/learners “can create coherent and cohesive text making full and appropriate use of a variety of organizational patterns and a wide range of cohesive devices.” At the C1 level, they “can produce clear, smoothly flowing, well-structured language, showing controlled use of organizational patterns, connectors and cohesive devices. Can produce well-organised, coherent text, using a variety of cohesive devices and organisational patterns.”

As it concerns the B2 level, users/learners: “can use a variety of linking expressions efficiently to mark clearly the relationships between ideas”, “can use a limited number of cohesive devices to link their utterances into clear, coherent discourse, though there may be some “jumpiness” in a long contribution”, “can produce text that is
generally well-organised and coherent, using a range of linking expressions and cohesive devices, and “can structure longer texts in clear, logical paragraphs.” Finally, at the B1 level, users/learners “can introduce a counter-argument in a simple discursive text (e.g. with “however”), “can link a series of shorter, discrete simple elements into a connected, linear sequence of points”, “can form longer sentences and link them together using a limited number of cohesive devices, e.g. in a story” and “can make simple, logical paragraph breaks in a longer text”.

3. Methodology and Corpus

Textual coherence should be investigated in extended discourse contexts. Due to their structure, corpora may provide all information required (Biber, Conrad & Reppen 2005). Our study is based on a learner corpus composed of writing data in French. Written composition activities were produced by candidates who participated in the differentiated and graded examinations of the National Foreign Language Exam System (Χριστιανικές Εκτελεστικές Περιοδοκίες Εκπαίδευσης, ΚΠΠΕ), which are organized twice a year under the auspices of the Hellenic Ministry of Education.1 The Greek state certificate for foreign language proficiency views linguistic pluralism as valuable to contemporary societies and offers exams in several European languages. The competences measured with the test papers are four: Reading Comprehension and Language Awareness, Writing and Written Mediation, Listening Comprehension, and Speaking and Oral Mediation.

The digitized corpus is enriched in an ongoing basis to maintain a balance in the representation of language proficiency levels (A1-C2), types of tasks, text types and communicative contexts to which candidates are asked to respond using the foreign language. Greek is the common language of all test takers. It is worth noting that the candidates taking these exams differ in their age group and literacy level. They also come from different regions and educational environments.

A criterion figuring among the evaluation criteria for the writing test is related to text grammar (text organization, coherence and cohesion). “Raters are trained to assess the degree to which candidates have managed to produce a coherent and cohesive script. Coherence refers to the presentation of ideas in a logical and understandable way. Candidates are expected to produce coherent texts by drawing on knowledge of how to organize and present their ideas from their previous experience as text producers and from their experience as readers.”2

In the present study, we use data gathered from written productions submitted for certification at levels B and C. We intend to investigate the variety in the organization and complexity of reference in discourse. As regards level B, candidates’ task was to: a) write an email to a friend to invite him/her to the 9th Festival of Plurilingualism (approximately 80 words), and b) write, and post on the website of Le Monde, an

1 https://rcel2.enl.uoa.gr/kpg/en_index.htm
article about the benefits citizens could enjoy from initiatives held by the municipalities of their country (approximately 100 words). At level C, candidates were asked to write an article which would be published on the site www.doctissimo.fr. This article should analyze the contribution of sports to the strengthening of social ties and explain how the public authorities could promote sports practice (approximately 350 words). The written outputs were expected to be informative and persuasive texts.

We used a sample of about 7000 words at levels B and C respectively. It must be noted that the corpus used was not annotated as regards anaphoras. This study was undertaken from a qualitative approach because elements used as anaphoric had to be separated manually after their automatic retrieval by means of a concordancer. Cases of exophoric elements (Martin 1992 : 122), e.g. definite articles, should be eliminated. Unitex was the tool we used to analyze our textual data. Unitex/GramLab is an open source, cross-platform, multilingual, lexicon and grammar-based corpus processing suite.

In the sections presented below, we turn our attention to the pronominal and the lexical anaphora as well as the cataphora.

4. Pronominal Anaphora

Pronouns are intended as a tool for the achievement of the referential continuity in discourse (Brown & Yule 1983). Several categories of pronouns may be used as anaphoric terms: personal, demonstrative, possessive, relative, indefinite, numeral. In both levels, personal pronouns predominate contrary to demonstrative pronouns. We noticed a quite significant presence of relative pronouns and the absence of possessive and numeral ones. Anaphoricity is not present only in the same sentence, it can also be observed in the context. As Ariel (1988 : 69) points out about the anaphoric references, “pronouns are predominantly used when the distances are short”. In many examples, the referent is mentioned in the first sentence and the pronouns, present in the sentences that come after, get their meaning from it. Anaphoric chains, that is succession of anaphoric elements, are not very common but there are examples where the pronouns used are the adequate ones.

Partial representation with en and sequences of clitic pronouns, e.g. le lui, la leur, have not been found in our sample corpus. The users/learners preferred to repeat the lexical elements used previously. Cases of zero anaphora were not observed, and this can be explained by the fact that they are probably unfamiliar with this characteristic of oral discourse. There were rare cases of omitting the complement of the verb needed to be represented by a pronoun:

1 https://unitexgramlab.org/
Il faut visiter le stade et participer dans ce programme parce qu’il est fantastique. Il ne faut pas négliger. [You must visit the stadium and participate in this program because it is fantastic. It should not be overlooked.]

Anaphoras belonging to the nominal category are the most frequent. Anaphoras with a verbal or an adjectival referent were not attested in the sample. This fact could be interpreted as a consequence of the difficulty faced by Greek-speaking learners to replace an adjective with *le* because of the absence of this possibility in Greek, e.g.: *Ils sont fiers d’eux. Ils le sont.* [They are proud of themselves. They are.]

What is needed to point out are the inappropriate and/or ungrammatical choices. If the learner makes a wrong decision because of choosing an anaphora that does not correspond to the referent, the understandability of the context is jeopardized. Ambiguous co-referencing in the text is in all likelihood due to the absence or the erroneous use of a pronoun. In the following text, the adjective *compétitive* agrees in gender and number with *personne*, whereas, immediately after that, *il* is used instead of *elle*. Similarly, in the case of *capacités*, there is no agreement in gender with the adjective *dernier*, but with the determiner *ces*, being common in both genders.

*Si une personne veut être compétitive au marché mondial, il doit acquérir des capacités. C’est vrai que la plupart des cours ne fournissent pas ces derniers.* [If a person wants to be competitive in the global market, he must acquire skills. It is true that most lessons do not provide these.]

In the following example, the referent of *ils* should be *les responsables*:

*Les volontaires seront encouragés par les responsables et ils les aideront à développer leurs connaissances.* [The volunteers will be encouraged by people in charge and they will help them to develop their knowledge.]

It is noteworthy that misunderstandings as regards the writer’s intention may be provoked by such faults.

The incorrect choice of *qui* or *que*, although it was found very rarely at level B, is another point to be noted which shows the insufficient acquisition of the relative pronouns.

*Ceci* and *cela* are two neutral demonstrative pronouns with a subtle difference in their usage. *Ça* is the informal version of the latter. *Ceci* is used to refer to something which shall be mentioned and is a cataphoric device. *Cela* refers backward to something that has already been mentioned. Users/learners of French cannot always distinguish the role of these elements and select randomly one of them. In our sample, we found *cela* 13 times at level C and once at level B. We found no occurrence of *ceci* in both levels. It could be reasonably assumed that, at level C, the knowledge about the colloquial

---

1 Examples translated
use of ça would have been acquired. Counterintuitively, the results are different. Ça\(^1\) occurred 3 times at level B and 15 times at level C, where it was used incorrectly. It must be noted that these productions received low assessment scores in all categories of evaluation criteria.

5. **Lexical Anaphora**

Anaphoric noun phrases were found in both levels. Definite articles, demonstrative and possessive determiners were used with an anaphoric functioning. Several types of associative anaphora (Kleiber 2001: 55-57) were found especially belonging to the type members-collection: famille-parents [family-parents], personnes-enfants [persons-children], loisirs-sport [hobbies-sport]. There were also observed sequences of functional associative anaphoras.

Faithful anaphora\(^2\), “a case in which the referent in not recategorized” (Willemse, Davidse & Heyvaert 2009 : 31), is present at both levels to represent several referents.

The unfamiliarity of users/learners with the importance of the interdependence of discourse elements not only inside the sentence limits but also in the entire text is plainly marked in cases like the one hereunder, where, apart from the deficient use of connectives and the unclear link between concepts concerning several capacités, the repetition of the noun personnes disorients the reader.

Les élèves d'aujourd'hui doivent acquérir des capacités certaines pour survivre dans notre monde globalisé. En d'autres termes, ils doivent avoir la capacité d'analyser des affaires internationales. C'est nécessaire de coopérer avec des personnes qui viennent d'autres civilisations. Les personnes qui vont acquérir ces capacités seront mieux équipées. [Students of today must acquire certain skills to survive in our globalized world. In other words, they must have the ability to analyze international affairs. It is necessary to cooperate with people who come from other civilizations. People who acquire these abilities will be better equipped.]

There were also occurrences of unfaithful anaphora, as it can be observed in the case of the hypernym sentiments or the synonyms participants and bénévoles:

Les élèves ont du stress et des angoisses. Au fil des années, ces sentiments peuvent devenir répressifs. [The students have stress and anxieties. Over the years, these feelings can become repressive.]

Les participants peuvent obtenir de l'expérience professionnelle. Ces bénévoles peuvent travailler dans plusieurs postes. [Participants can gain work experience. These volunteers can work in several positions.]

---

1 «Lorsque /ça/ renvoie à la situation, il est « déictique » ; quand il réfère au discours, il est anaphorique.» (Maillard 1974 : 66)

2 *Anaphore fidèle* and *anaphore infidèle* (Maingueneau 2000 : 175) were translated as *faithful anaphora* and *unfaithful anaphora* (Willemse, Davidse & Heyvaert 2009 : 31).
In the example set out hereafter, *ce besoin* probably replaces two word sequences: *développement des compétences globales* and *naissance des compétences et des capacités*:

*En ce qui concerne les raisons du développement des compétences globales, nombreuses raisons ont contribué à ce besoin. Le monde entier sait la problématique qui a conduit à la naissance des compétences et des capacités. Ce besoin a...* [Regarding the reasons for the development of global competencies, many reasons have contributed to this need. The whole world knows the problem that led to the birth of skills and capacities. This need has...]

Cases where the anaphora is not a pronoun but a noun phrase were noted, e.g. in the following example *cet âge* is the anaphora referring to the referent *quinze ans*:

*Salut ! Je m'appelle Martha et j'ai quinze ans. Tous les adolescents dans cet âge ont des problèmes.* [Hi! My name is Martha and I am fifteen years old. All teenagers in this age have problems.]

To reformulate or explain the referent, an important variety of noun phrases *ce + general noun* was used in the case of action taken by the municipalities, previously mentioned in the productions. The relation of co-reference is based on an attribution of this property to the action taken, made implicitly. *Cette mesure, cette pratique, cette modernité, cette solution, cette façon, cette façon de vie, cette méthode, ce phénomène, ce changement* are anaphoras used. Sometimes, this device was used twice:

*En ce qui me concerne, cette pratique n'est pas correcte. Je suis contre ce phénomène.* [As far as I am concerned, this practice is not correct. I am against this phenomenon.]

In most cases, a referent corresponds to a noun or an adjectival phrase. Nevertheless, there are cases of resumptive anaphora where the referent corresponds to a larger segment of the sentence.

Repetition is essential in some cases, but, sometimes, a word hovers around the text without having any objective and the use of a pronominal element or a determiner would be the most appropriate option. An important observation we often made in both levels concerns the iteration of a key word throughout the text. This approach may be related to the fact that pronouns eventually provoke ambiguities in discourse, or the candidates may have chosen a surest route against the risk of using an incorrect pronoun.

*Le musée organise beaucoup d'activités pour les enfants qui vont avoir lieu dans ce musée.* [The museum organizes a lot of activities for children that are going to take place in this museum.]

In the following example, the referent is the same in almost the whole production reducing the possibility of the existence of other referents.
Je voudrais parler des bienfaits de la musique. On doit aimer la musique. Il y a une organisation qu'on peut faire des activités avec la musique. On doit écouter à la musique parce que la musique est très bonne pour la santé. [I would like to talk about the benefits of music. We must love music. There is an organization that we can do activities with music. One should listen to music because music is very good for health.]

Although anaphoric elements are used adequately in a text, sometimes, in the same text, the lexical anaphora that is introduced by ce, cette, ces is distant from the referent and it might be difficult for the reader to resolve the reference, e.g.:

Un programme de volontariat est une bonne solution. On proposera une façon contemporaine pour devenir un volontaire qui va participer pour aider les autres personnes. Les volontaires seront aidés par les responsables. On peut donner la chance aux participants de gagner de l'expérience et des savoirs. Ainsi, les volontaires, après leur enseignement, seront capables de travailler dans nombreux domaines d'action intéressants. Cette initiative est... [A volunteer program is a good solution. We will propose a contemporary way to become a volunteer who will participate to help other people. The volunteers will be helped by the leaders. Participants can be given the chance to gain experience and knowledge. Thus, the volunteers, after their education, will be able to work in many interesting fields of action. This initiative is...]

Cette initiative is used here in place of Un programme de volontariat. The two noun phrases are separated by four sentences.

Adverbial anaphora was found five times at level B and four times at level C. Regarding on adjectival anaphora, no occurrences were observed at level B but it was found eight times at level C with the structure tel + noun. There were also cases of indirect anaphora (Gardelle & Vincent-Durroux 2022). In the following text, there is no referent. The lexical unit aussi helps us to understand, after a sequence of complex inferences, that Greece is not the only country where innovation in the domain of sports is welcome:

Des programmes d’innovation et de motivation sportive doivent exister en Grèce aussi. [Sports innovation and motivation programs must exist in Greece too.]

6. Cataphora

Speaking about the entire sample corpus, cataphoric mechanisms are not abundantly used. Nevertheless, there are interesting cases that attracted our attention. In our corpus, cataphora is either interphrastic or transphrastic (Kęsik 1989 : 36-43). Different cases which show how cataphora is used are shown below:

Pour son développement personnel, la jeune personne doit acquérir des connaissances nécessaires. [For his personal development, the young person must acquire the necessary knowledge.]
Les actions des participants sont nombreuses. Ils peuvent participer en programmes sociaux ou à l’organisation de rencontres. [The actions of the participants are numerous. They can participate in social programs or in the organization of meetings.]

Je suis une adepte du premier choix, c’est-à-dire je préfère de rester en classe un peu plus longtemps. [I am a fan of the first choice, that is to say, I prefer to stay in class a little longer.]

Les jeunes peuvent aider avec un support qu’ils utilisent au quotidien : l’ordinateur. [Young people can help with a medium they use daily: the computer.]

In the same way as anaphora, cataphora can be segmental or resumptive. In our sample corpus, we found a few examples of resumptive cataphora. In relation to examples of segmental cataphora, none was found.

It would also be interesting to note the ana-cataphoric use of the pronoun le in the following example. le replaces le programme sportif of the previous sentence and announces the subsequent discourse (Kęsik 1989: 79).

J’ai trouvé des renseignements pour le programme sportif. Laissez-moi le présenter. D’abord, à Athènes, il y a des volontaires... [I found information for the sports programme. Let me introduce it. First, in Athens, there are volunteers...]

Solution was used several times in the first written production of level B. In the first two examples, it is used successfully in a cataphor where the explication follows immediately after, and, awkwardly, in the third example, where the sign of equality (=) interferes in the text:

J’étais triste et j’ai décidé de discuter avec mes parents pour une meilleure solution. On a trouvé un professeur et j’ai commencé des cours. [I was sad, and I decided to discuss with my parents for a better solution. We found a teacher and I started lessons.]

Je te donne deux solutions à ton problème : tout d’abord, [...]. Ensuite, [...]. [I give you two solutions to your problem: first, [...]. Next, [...]]

Ma solution pour ton problème est = il faut faire plus attention en classe. [My solution for your problem is = you must be more careful in class.]

There is also a case where the same noun is used as a cataphor at the beginning of the text and as a conceptual anaphor at the end of the text:

Je lis l’e-mail et j’ai les solutions pour ton problème. [...] J’attends ta réponse pour ces solutions. [I read the email and I have the solutions for your problem. [...] I await your response for these solutions.]

7. Conclusions and suggestions for future research

This paper encompassed the question of the presence of the anaphora and cataphora in a sample corpus composed of texts of the language proficiency levels B and C of the

The findings of this modest sample size study led to the conclusions outlined below:

a) Anaphora is an essential element in the written production of the candidates, and it accounts for the predominant part of endophora. Cataphoric reference examples were also found in our data but to a lesser extent.

b) Both pronoun anaphora and lexical anaphora are present in the text. Anaphoras belonging to the nominal category are the most frequent. On the other hand, anaphoras with an adjectival or a verbal referent were not attested in the sample. Reference was also made to the use of noun phrases ce + noun and the replacement of the referent by another noun. Faithful and unfaithful anaphora are also used.

c) It should be noted that the linguistic errors committed by the candidates were not considered in the framework of this study. However, errors regarding the reference were noted because they usually impede the comprehension. It is noteworthy that restructuring of sentences is in some cases considered necessary. The distance between the referent and the anaphora and, also, the context which stands between them may sometimes lead to confusion.

d) The redundant use of the referent throughout the text is also an observation made. Its replacement by the most appropriate pronoun, or any linguistic element, but also the information relating to the risk of overuse of this element should be considered by the teacher as a challenge.

e) Regarding the presence of anaphors and cataphors in the sub-corpora of levels B and C, the research revealed no significant differences between them except some cases, e.g. the use of cela or the adjectival anaphora). More indications should be necessary to confirm or not that users/learners do not behave differently.

French is a language which disposes a variety of anaphoric devices that they are not always introduced or considered as such during the teaching process. Teachers should identify the characteristics of anaphora and explain the role of all cohesive devices found in texts exploited in the classroom of the foreign language for the enhancing of comprehension skills. Corpora are a precious instrument to provide evidence on text elaboration. They can be used to facilitate learning and to assist learners to reflect on the correct interpretation of relationships among elements.

Writing a cohesive text by connecting ideas in an appropriate way is an important skill for a successful communicative process, but it is not the only one that should be fostered by teachers. It should be noted that proofreading of learners’ written productions by themselves is a very important skill that should be tackled with particular care in the classroom of the foreign language. It is an opportunity for them to provide for the quality of their productions and to develop their self-awareness. On
the other hand, teachers should show professional awareness to develop this capacity to learners. They should enable learners to manage their time so that they can have the opportunity to read very carefully their completed texts before submission in order to reveal inconsistencies, ambiguities and lack of clarity. The time devoted to proofreading should be considered as an essential part of the writing procedure and not as a luxury or an unnecessary final step. As it happens in real life, in the professional and social domains, proofreading strategies should be also appreciated and prevalent in the teaching of foreign languages.

Regarding the generalizability of the conclusions in our limited study, it should be noted that the combination of quantitative and qualitative research could probe into the issue we are dealing with and offer reliable results. To achieve a more complete description in the hope that it will illustrate what users/learners can do or cannot do, it would be useful to verify the conclusions drawn here in a large-scale, more representative, dataset compiled from written productions of varying text type or genre and a range of topics treated. Argumentative or narrative strategies keep abreast of the relationships among the different elements in context. Data from users/learners’ not having a common mother tongue could also be explored in order to find out if this is a significant factor as regards the use of anaphora. Furthermore, the annotation of anaphoric reference in learners’ corpora could be a useful tool for the extraction of statistical data and the exploration of this domain in users/learners’ interlanguage. It should also be useful to explore the appropriate use of cohesive ties in written productions having a good evaluation in the other criteria imposed by the evaluation grid of the examination system.

The compilation of the corpus used for this research was supported by the Research Committee of the Aristotle University of Thessaloniki.

References


