CORPUS LINGUISTICS AND METAOPERATIONAL GRAMMAR TO THE TEACHING OF GRAMMATICAL COHESION

Kouakou Yannick KONDRO
Assistant, Enseignant-Chercheur
Université Félix Houphouët-Boigny
yannickkondro@yahoo.fr

Abstract: Producing cohesive writing can be a difficult task if foreign language learners are not presented with the necessary analytic tools to fathom out the functioning of language. Studies show that cohesive writing is particularly difficult for non-native English speakers (E. Hinkel, 2001, 2002; J.M. Swales & C.B. Feak, 2004). And it is also proved that textual cohesion devices do not receive enough attention in traditional language teaching (G. Cook,1989 p.127). That is undoubtedly the reason why there is a shortage of teaching materials on cohesion (M. Johnson, 2017). Consequently, this analysis aims to propose an approach that can be helpful to English learners in their writing. It also ambitions to familiarize students and teachers with online corpora in addressing some linguistics facts. It draws on metaoperational grammar theory's findings (H. Adamczewski, 1982, 1991; J-P. Gabilan, 2011; Y.K.J-F. Kpli, 2014). Utterances are extracted from online sources, like Corpus of Contemporary American English (COCA) and British National Corpus (BNC) in order to take advantage of this revolutionary approach to data collection (Corpus Linguistics), not yet widely applied in Côte d'Ivoire. It results from the analyses that metaoperational grammar theory and corpus linguistics can help students to produce cohesive writing in order to succeed in conveying their own ideas.

Keywords: grammatical cohesion, metaoperational, grammar, corpus linguistics

LINGUISTIQUE DE CORPUS ET GRAMMAIRE MÉTAOPÉRATIONNELLE À L'ENSEIGNEMENT DE LA COHÉSION GRAMMATICALE

Résumé : Écriture de manière cohérente peut être une tâche difficile si les apprenants des langues étrangères ne disposent pas d'outils d'analyses nécessaires pour comprendre le fonctionnement de la langue. Les recherches montrent qu'écrire de manière cohérente est particulièrement difficile pour les apprenants de la langue anglaise (E. Hinkel, 2001, 2002 ; J.M. Swales & C.B. Feak, 2004). De plus, il est également prouvé que les outils linguistiques qui permettent de produire la cohésion textuelle ne reçoivent pas suffisamment d'attention dans l'enseignement traditionnel des langues (G. Cook, 1989 p.127). C'est, sans doute, la raison pour laquelle il existe un besoin sur les supports pédagogiques sur la cohésion (M. Johnson, 2017). Par conséquent, cette analyse vise à proposer une approche qui peut être utile aux apprenants d'anglais dans leur rédaction. Elle se donne également pour objectif de familiariser les enseignants et apprenants à la linguistique de corpus dans l'analyse des faits linguistiques. Elle s'appuie sur les résultats de la théorie de la grammaire métaopérationnelle (H. Adamczewski, 1982, 1991 ; J-P. Gabilan, 2011 ; Y.K.J-F. Kpli, 2014). Les énoncés sont extraits de sources en ligne, comme le Corpus of Contemporary American English (COCA) et le British National Corpus (BNC) afin de tirer profit de cette approche révolutionnaire de la collecte de données (Corpus Linguistics), encore peu appliquée en Côte d'Ivoire. Il ressort des analyses que la théorie de la grammaire métaopérationnelle et la linguistique de corpus peuvent aider les élèves à produire des textes cohérents afin de réussir à véhiculer leurs propres idées.

Mots-clés : cohésion grammaticale, métaopérationnelle, grammaire, linguistique de corpus

Introduction

Writing presents challenges for many learners of the English language. Studies show that cohesive writing is particularly difficult for non-native English speakers (E. Hinkel, 2001, 2002; J.M. Swales & C.B. Feak, 2004). Moreover, it is proved that textual cohesion devices do not receive enough attention in traditional language teaching (G. Cook,1989 p.127). That is undoubtedly the reason why there is a shortage of teaching materials on cohesion (M. Johnson, 2017). Thus, this analysis aims, first and foremost, to give the necessary analytic tools to English learners in order to reach cohesive writing. Secondly, it provides teachers with the know-how to present cohesive operators, especially the grammatical references. And finally, it ambitions to familiarize students and teachers with online corpora in addressing some linguistics facts. Two main approaches to the analyses are provided, namely corpus linguistics and metaoperational grammar. When collecting real data in the analysis of linguistic facts came into its own, linguists had to read lots of books, newspaper etc. and listen to a great range of real communication in the hope of finding the linguistic facts to single out. Today, the advent of computer allows linguists to have access to a large scale of data. It is now possible to search rapidly and reliably through millions of words from databanks (for example, BNC and COCA). This revolutionary approach to data collection through the computer in the study of languages is known as corpus linguistics. Corpus linguistics is undoubtedly one of the "fastest-growing methodologies[1] in contemporary linguistics" (S. T. Gries, 2009). Unfortunately, while it "presents interesting ways and possibilities" (A. M. Pantaleon, 2018 p.151) in teaching and analyzing languages, its use remains rare in the educational institutions in Côte d'Ivoire. Thus, this analysis also aims to draw on corpus linguistics in order to help researchers in the field of metaoperational grammar take advantage of it. As far as metaoperational grammar (H. Adamczewski, 1982, 1991; J-P. Gabilan, 2011; Y.K.J-F. Kpli, 2014) is concerned, it is applied to decode the functioning of the cohesive devices in order to show how they can be used into discourse. The "descriptive and explanatory power" (H. Adamczewski and J.P Gabilan, 1992) of this theory can help students to acquire what they need to know in order to succeed in conveying their ideas.

1. The Rationale of the Study

Cohesion is one of the greatest qualities in writing skills. In their seminal work on *English cohesion*, M.A.K. Halliday and R. Hasan (1976: 299) note that cohesion "cohesion expresses the continuity that exists between one part of the text and another". It is "the property of connectedness that characterizes a text in contrast to a mere sequence of words" (M. Mahlberg, 2006: 363). Considering its significance in students' productions, it is essential to reflect on methods to bring solutions to the issue. Indeed, cohesion is all-important because it helps understand a text. Students know how sentences or paragraphs are related. It also contributes to the readability of a text. An English learner can develop a native-like oral and written competence if he or she connects properly sequence of words. Cohesion also helps students write cohesively because the texts that are not cohesively

---

[1] Whether it a methodology or a theory is still an ongoing debate. In this article, we side with the idea that it a methodology applicable to Metaoperational Grammar.

written create ambiguity to the readers. Two main reasons can be found in conducting this analysis. The first one is to prove that metaoperational grammar is a linguistic tool that can help teach cohesion. In fact, metaoperational grammar findings can be used to explain and teach about cohesion, especially grammatical cohesion. In that perspective, we put into practice the recommendation of H. Adamczewski (2004) who stresses out that English linguistic theory should not be far from the practice. In other words, he believes that linguistics theories should lend themselves to the practice. In his speech of May 22nd, 2004, he highlighted: "La recherche en linguistique anglaise n'est pas du domaine de l'art pour l'art, […] La recherche théorique doit trouver son application dans le geste didactique.". And the theory in use in this study allows the linguist to practically apply its findings into teachings. In the analyses, we have recourse to the underlying findings of grammatical reference forms like "this"; "that", "A[2]", "the" and the pronouns to show how they are all expressed and can be used by learners. The second reason is that corpus linguistic can be of great help in teaching cohesion. Far from trawling through books and listening to a great number of podcasts in order to identify the linguistic operations aimed to be proved, this analysis shows that we can take advantage of electronic corpus linguistics to have millions of utterances at one's click. As a matter of fact, the most obvious advantage is the speed in which one can get to data. The corpora offer difference type of sources with the context from which they are extracted.

2. Methodology

Two main methodologies are to be noted in this analysis. The analysis consists in explaining the use of some linguistic devices in order to show how they can be used in academic writing. In the analyzing process, the linguistic units are contrasted with other units close in the system of language, that is, units that are interchangeable, i.e., "this" and "that". The grammatical units under study are known as reference forms. Their meanings are not readily identifiable, therefore what can discriminate them is their metalinguistic roles. These particular nuances in their functioning help produce substantial differences in meaning construction. As a matter of fact, they should not be overlooked or left out in meaning encoding or decoding. This approach to unit analyses is contrastive. In fact, contrastivity is a methodological prerequisite in this theory. The contrastive approach applied in this analysis is the intralingual contrastive one. In practice, some units are substituted and commuted in order to show how utterances can be cohesive or not. We then explain, in the light of the theory in use, why some sentences are cohesive or not. With that in mind, our approach is descriptive and explanatory. As far as the collection of data is concerned, metaoperational grammar lays a strong emphasis on analyzing authentic data collected. That is the reason why, the utterances are taken from online corpora, i.e., Corpus of Contemporary American English (COCA) and British National Corpus (BNC). Those corpora are the only large, well-balanced corpora of English that are freely available online. The following table gives an overview of both corpora:

---

[2] In the presentation of reference forms in M.A.K Halliday and R. Hasan (1976), "A" is not counted. The reference to it is to account for the use of "the" because in the analysis of metaoperational grammar, both "A" and "the" constitute a microsystem.

Table 1: Comparison of Number of Word between COCA and BNC

COCA	BNC
The corpus contains more than **one billion** words of text (25+ million words each year 1990-2019) from eight genres: spoken, fiction, popular magazines, newspapers, academic texts, and (with the update in March 2020): TV and Movies subtitles, blogs, and other web pages.	The British National Corpus (BNC) contains 100 million words of text from a wide range of genres (e.g. spoken, fiction, magazines, newspapers, and academic).

The relevance of the corpora is that natural utterances exist and can be used to explain why people use some units instead of others. It is, therefore, not about creating our own sentences to suit the analyses' goal.

3. Grammatical Cohesion

M.A.K. Halliday and R. Hasan (1976) identified two types of cohesion, namely grammatical and lexical cohesions. Grammatical cohesion includes reference, substitution, ellipsis and conjunction. This paper focuses on the grammatical referencing because it constitutes a pillar that contributes to make the texture of the text well-connected. Grammatical reference is used to refer to linguistic items in order to create cohesion in a piece of writing. It is "a set of grammatical resources that allows the speaker to indicate whether something is being repeated from somewhere earlier in the text" (N. Chanyoo, 2018: 995) Its purpose is to avoid repetition and redundancy. M.A.K. Halliday and R. Hasan (1976: 31) indicate that referential forms are "forms which instead of being interpreted semantically, in their own right, […] make reference to something else for their interpretation". There are two types of references, namely exophoric and endophoric references. When the interpretation lies within a text, they are called endophoric reference. This sub-type is composed of two different types: the anaphoric and cataphoric references. The anaphoric reference occurs when a word or phrase refers to something mentioned **earlier** in the discourse. For example, the following utterance, "he" occurs as a substitute to "Nigel":

(1)

> Nigel is so laid-back and a real joker but **he** works like hell and gets the horses very fit.' I think **he** is brilliant. **He** makes training look easy and is a man with a terrific future. BNC

The cataphoric reference occurs when a word or phrase refers to something mentioned later in the discourse. For example, "this" presents the segment: "Why? What, at this juncture, does it matter?" in the following utterance:

(2)

> The message in the doctor's silence was **this**: " Why? What, at this juncture, does it matter? COCA

When the interpretation is extra-textual, that is, when it lies outside of the text of the discourse, in the situation of communication, the relationship is said to be an exophoric. In effect, M. Lederer (2014) notes that

> La situation représente tous les éléments de la perception sensorielles non linguistiques concomitants au discours... la situation est le cadre matériel, la salle où l'on se trouve, la vue que l'on a, les gestes et les mimiques de l'orateur, tous les éléments formellement présents et perçus en même temps que le discours; lorsqu'elle n'est pas pertinente, la situation peut parfois distraire l'attention; Lorsqu'elle est, elle oriente la compréhension de l'énoncé linguistique vers le sens voulu par le locuteur
>
> M. Lederer (2014:48)

The extra-textual reference can be difficult to identity in the corpus linguistics where only texts are found, unless there is a specification in the utterances under study. The differences established above can be illustrated by the following diagram:
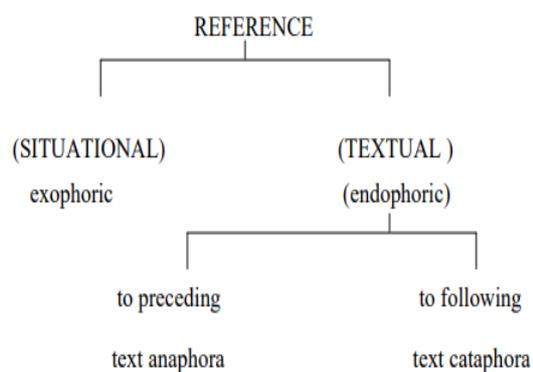
M.A.K. Halliday and R. Hasan (1976, p. 33)

It should be noted that functionally speaking, there are three main types of cohesive references, namely the personal pronouns, the demonstratives and the comparatives. In the demonstratives subgroup, the elements in the micro-systems "this/that" and "here/there" are set different in terms of distance. "This" and "here" are said to indicate proximity to the speaker while "that" and "there" express remoteness. Though proximity can be one of the criteria, it does not help to account for all the occurrences of those units. To prove this point, H. Adamczewski and J.-P Gabilan (1992, p.117) present different utterances, amongst which the following: (Context: the speaker is at a desk surrounded by all the inhabitants of the house where a crime has just been committed. There is a scratch on the desk.)
(3)
        Hullo! Did you see that scratch?

In this context, the explanation referring to the distance is not justified. In the perspective of Metaoperational Grammar, those units are the result of the principle of phases addressed under 7. Besides, how far can something be said to be far? Why is "this" used instead of "that" in utterance (4)?

(4)

       - Is there life on other planets?

       - On those nearest to us, probably not. (H. Adamczewski and C. Delmas, 1982: 221)

"Those" (plural of "that") implies that the planet is far. Why then do we have "nearest" and not "farthest"? These utterances reveal that the question of distance in particular, and the reference to the extra-linguistic domain to explain grammatical facts (intra-linguistic domain) in general render the explanations of grammatical facts difficult to understand. That is the reason why our approach to grammatical cohesion is meta-linguistic oriented.

4. Corpus Linguistics

Corpus Linguistics is the study of linguistic phenomena through large collections of computer-readable texts. This approach is defined by R. Nordquist (2019) as the "study of language based on large collections of 'real life' language use stored in corpora—computerized databases created for linguistic research." A corpus (plural: corpora) is a body of naturally occurring language, that is to say the language of the corpus is above all real, based on actual usage. Strictly speaking, this form of corpora can be referred to as computer corpora because the collections of naturally occurring examples are computer-based. Though there are different ways to define a corpus (T. McEnery *et al.*, 2006), it presents two key qualities: a corpus is machine-readable and authentic, that is, it should be based on real, rather than contrived, data (C. Meyer, 2004). That form of collecting data dates back to the early 1960s when the first corpus, the Brown Corpus (N. Francis and H. Kučera), was created. The corpus consisted of one million words of American English texts printed in 1961. At that time, it was not warmly accepted amongst the members of the linguistic community because generative grammarian[3],which was the dominant theory, upheld the idea that "the only legitimate source of grammatical knowledge" about a language was the intuitions of the native speaker, which could not be obtained from a corpus" (C. Meyer, 2004:1). It was then so severely criticized that it became marginalized because of the alleged 'skewedness" of corpora. But with the development of technology offering massive storage and increasing processing power, the exploitation of massive corpora is achievable. With this approach to data collecting, corpora can be objective and representative (T. McEnery *et al.*, 2006). Before the development of computer corpus linguistics, we had a form of data collection that can be traced back to Otto Jespersen's multi-volume *A Modern English Grammar on Historical Principles* (1909–49). In fact, his tremendous work was based on "a corpus representing the canon of English literature: thousands of examples drawn from the works of authors such as Chaucer, Shakespeare, Swift, and Austin that Jespersen used to illustrate the various linguistic structures he discusses" (C. Meyer 2004: xii).

---

[3] It is to be noted that in the 1960s, generative grammar dominated the field of linguistics and ideas that were not aligned to what generative grammarian deemed acceptable had little chance to thrive.

5. Traditional Approaches to Data Collection

In the teaching and analysis of languages, two trends dominate the data analysis approach. One trend is intuition-based. In this case, the researchers, mainly the grammarians, create sentences based on their own intuitions to explain language phenomena. Since intuition is readily available, the researchers can invent examples instantly. The problem with those invented sentences is that they lack context of production. T. McEnery et al. (2006) highlight two reasons why intuition should be applied with caution

> […] first, it is possible to be influenced by one's dialect or sociolect; what appears unacceptable to one speaker may be perfectly felicitous to another … Second, when one invents an example to support or disprove an argument, one is consciously monitoring one's language production. Therefore, even if one's intuition is correct, the utterance may not represent typical language use.
>
> T. McEnery et al. (2006: 6)

Intuition cannot explain or cover all aspects of a language. Consequently, there is a high possibility to have linguistic explanation that are riddled with exceptions because, in that case, it is quite difficult for the linguist to propose a generalizable rule. Multiple exceptions to grammar rules are proofs that there is a need to analyze the linguistic issue developed further (Y.K.J.F. Kpli, 1992). J.P. Gabilan (2016) highlights that "rules are meant to be rules and rules tolerate no exceptions". A case in point is the explanation of the functioning of "be+ing" (H. Adamczewski, 1976, 1982). It was the cause of a number of research because findings on the matter always met with exceptions. It could be read in almost all the grammar books that "be+ing" (also known as the progressive) "cannot be used" with some verbs – e.g.: love, agree, mean etc. But English-speaking people break this rule on a daily basis in real communication settings. The clicks on the electronic corpora show the following examples:

(5)

Auntie El's bloods were still good yesterday so she's still on schedule. She's going to have a PET scan after her 4th treatment. We're **hoping** her bloods stay good and she stays on track! BNC

(6)

Now I like old people as much as the next person. I'll be one soon myself. So, I'm not **hating** old people here. BNC

(7)

I don't think a lot of people are **believing** you're making millions of dollars at nine. BNC

The work of H. Adamczewski is groundbreaking because the findings are generalizable. With actual utterances used on a daily basis, he comes to the conclusion that analyzing language free from language-external influences existing in the natural occurrences is counter-productive. H. Adamczewski also proves that the direct assignation of meaning to meaningless linguistic units in general, and to the so-called progressive form, is the main reason for the failure of the descriptive approach to account for the working of language. Another intuition-oriented process to get actual utterances is to ask native speakers to approve of a use or not. The native speaker is considered as an informant. Though, it is

correct and can help in languages that are not yet written, generalizing claims about spoken English, requires a suitable spoken dataset because "the speech of one person alone is unlikely to provide a suitable basis for such generalizations." (T. McEnery and A. Hardie, 2012 p.6). Besides, "results based on introspections alone are difficult as introspection is not observable." (T. McEnery *et al*. 2006: 6) . The second trend is real-life corpus-based approach. When the idea of collecting real-life data came into its own, it was painstaking to get data. The linguist was to trawl through a considerable number of texts in the hope of finding the examples that would help him explain a phenomenon. This task was demanding and strenuous because the analyst had to read and listen to real communication in order to extract the utterances that would justify his/her analyses. Though, it is a rewarding approach, the written or transcribed texts were not representative for any particular language.

6. Metaoperational Grammar

Metaoperational Theory was founded by H. Adamczewski (1982) and further enriched by *les Amis du Crelingua*[4]. It is developed on the assumption that utterances exhibit visible traces of the utterer's underlying (invisible) structuring process. H. Adamczewski (1982, p.5) puts: "the linear word sequence contains visible traces of invisible operation[5]". Since the invisible structuring process is not directly reachable, there is a need to look into the surface manifestation of language before reaching the invisible. To unveil the underlying functioning, a clear distinction is made between the notion of sentence and that of utterance. He then chooses to work on utterances because they are taken out of natural contexts of communication, while sentences can be seen as the invention of an analyst based solely on his/her intuition. The theory is, therefore, focused on collecting data from different sources, oral or written, provided they come from real communication settings. From that perspective, it is initially a corpus-based theory. It then stood as a new way to approach language compared to the dominating trends at its birth. The difference between this approach to language and the classical approach is put forth in the following table:

---

[4] The objective of the association "Les Amis du Crelingua" is to promote research in linguistics and to highlight the operating principles that govern the functioning of languages.
[5] The original quotation is: "les énoncés de surface comportent des traces visibles d'un fonctionnement invisible".

Table 1. A University Grammar of English vs. Grammaire Linguistique de L'Anglais

	A University Grammar of English	Grammaire Linguistique de l'Anglais
1	descriptive / structuralist	explicative / post-structuralist
2	monolingual approach	intra- and interlingual approach
3	direct assignment of meaning to formal categories / referential function	metalinguistic function of operators
4	context insensitive	context- and corpus-based analysis
5	linear description	"The linear is not the message."
6	centrality of the grammatical subject	centrality of the linguistic subject
7	sentences / language	utterances / discourse
8	focus on macro-structures	focus on grammatical micro-systems
9	taxonomic and rule-based	core-value approach / invariance theory

M.-H. Kahlaoui (2019: 5)

F. Lachaux summaries the general objective and the method apply in metaoperational grammar in these terms:

> L'approche méta-opérationnelle, élaborée par H. Adamczewski, propose d'aborder les problèmes complexes que pose l'interprétation des marqueurs linguistiques, traces visibles du travail de structuration effectué en amont de l'énoncé, en allant décrypter, en-deçà de l'ordre linéaire de surface, l'ordre des "étapes de la production réelle de l'énoncé" que ces marqueurs "réfléchissent" et permettent de "chiffrer", et ce dans une démarche contrastive, car les langues "s'entre-éclairent" et laissent transparaître les mêmes opérations profondes à l'aide de marqueurs différents.

F. Lachaux (2006:38)

It would be pretentious to aim to cover all the principles developed in this theory. For the sake of this analysis, we will focus on two of them, i.e., the invariant value and the phase systems. The invariant value is the kernel value of a given linguistic unit. It is, in fact, the common denominator linking all the speech effects (uses) of a linguistic unit. More often than not, we stuck with the idea that some words do have absolutely the same role or meaning, forgetting that language is a system. That is the reason why it is important to have the tools that help set words apart by getting into their intimate functioning. Thus, advocates of this theory track down the invariant value to understand the particularities of linguistic units. The invariant value is a bedrock that can help justify that since language is a system, no two units can have the same meaning. The system of phases posits that most grammatical phenomena are organized systematically in pairs, known as phase 1 and phase 2. According to Adamczewski, this basic principle is repeated cyclically to create different grammatical tools that are necessary to the working of languages. The phase 1, also known as the assertive status (J.P. Gabilan, 2011), is not presupposing while the phase 2 is. The following examples elaborate on the phase system:

(8)

> I've got **a new car** and it's already got 23,000 on the clock, and three-quarters of that must be to the club. BNC

(9)

> We played golf together, then he took me for a ride in **the new car**. BNC

In utterance (8), the co-speaker does not know anything about the car. The speaker is informing him that he got a new car. In this case, (A) is said to be phase 1 or assertive because it is used to give fresh information, to establish fresh relation. But, in utterance (9), the speaker is not informing the co-speaker about the car. The existence of "car" is a shared knowledge, that is, both of them share the information about the existence of the car. "The" is used to recall "car". "The" is, therefore, phase 2 or non-assertive.

7. Cohesion in the Light of Metaoperational Grammar

In the following analyses, we show how using the system of phases can help develop strong cohesion in writing. As stated earlier, the grammatical references are analyzed in order to show how they can cohesively be used in students writing. Therefore, the utterances analyzed contain the operators "a"; "the"; "this"; and "that" and the personal pronouns. The following analyses will help better present the functioning of each of these operations in order to reach cohesive writing.

(10)

> I've given you everything he gave me.
> Are you sure of that, sir? COCA

In this utterance, the occurrence "this" would be impossible because it would imply that the speaker is about to present something to the co-speaker. In fact, the speaker is referring to the previous segment already introduced by the co-speaker. The segment "are you sure of that" can be paraphrased by "are you sure of what you have *just said*? Asking the question with "sure" entails that the co-speaker is aware of the issue, for we cannot ask a person to confirm something that he/she has no information about. The use of "that" in that context is the trace of the underlying functioning of "that": it is used to refer back to something already used into discourse. The following utterances show how important is the invariant value.

(11)

> **A person** is not eligible for an elective municipal office unless **that person** is a qualified elector of the municipality and has resided in the municipality at least one year next preceding the election or appointment. COCA

(12)

> As for Friedman's first two ways to spend money, they both involve **a person. That person**, having a heart, recognizes a need and uses his money to meet it. The interests are aligned. COCA

Utterance (11) and (12) are examples of the cohesion expressed by the speaker in his structuring. The introduction of "A person" make it impossible for the speaker to use an

operator of phase 1 if he or she wants to imply that the segment being introduced has already been introduced into discourse. Any reference to a segment already used call for a phase 2 operator. It would sound awkward to read:

> ***A person** is not eligible for an elective municipal office unless **a person** is a qualified elector of the municipality…
> ***As for Friedman's first two ways to spend money, they both involve a person. That person,** having a heart…

The awkwardness is due to the fact that "a" is used to establish a fresh relation with the noun. In other words, it is used to introduce the noun for the first time into discourse. That is the reason why "that" is used. The speaker implicitly aims at showing that the second use of "person" is the reintroduction of the prior use of "person". The use of "this" and "that" can be problematic, particularly for French-speaking people, because there is only one operator in French to translate them. The French operator "*ce*" is their counterpart. In this sense, the translation from English to French is straightforward whilst the other way around requires some metalinguistic operations not readily identifiable. In order to provide cohesive translation in this case, there is a need to understand the metalinguistic role to "*ce*" to match it with that of "this" or "that" (This analysis will require another article). Let us consider the utterance:

(13)

> The 1969 film Z, which satirizes the military dictatorship ruling Greece at that time, has **this notice**: " Any resemblance to actual events, to persons living or dead, is not the result of chance." COCA

In (13), the speaker aims to present the notice "Any resemblance to actual events, to persons living or dead, is not the result of chance" to the co-speaker. He assumes that this information in not shared by the co-speaker. "That" could occur with or without that following segment. If it occurs without the segment, the speaker assumes that the co-speaker is perfectly aware of that notice. But if "that" occurs with that segment, it means that the speaker estimates that the co-speaker has already the information, but he may have forgotten, therefore there is a need to recall him with the full information.

The operator "the" has the same metalinguistic role as "that". "The" is used to recall a piece of information. The utterances below disclose their internal functioning.

(14)

> It was dug up two years later. They put it in the back of **a car. The car** was squished down to a four-by-four foot square and sold as scrap metal with several thousand other cars, which is shipped over to Japan, and that's what he said happened to Jimmy. COCA

(15)

> SECURITY chiefs and the SAS last night lashed **a new book** they described as a bomber's guide to Buckingham Palace. **The book** includes precise plans pinpointing the bedrooms of the Queen, Prince Philip, Prince Andrew and Prince Edward. BNC

"A" is used to introduce "car" and "book" into discourse. When the speaker comes back on the segment introduced, he uses the phase 2 operator "the". The pronouns do have the same metalinguistic roles, as can substantiated in the utterances:

(16)

> **Anne Midgette** came to The Washington Post in 2008, when **she** consolidated her various cultural interests under the single title of chief classical music critic. **She** can be found online as The Classical Beat. COCA

(17)

> Oh, no. I was thinking of my **mother. She** can be so possessive. COCA

"Anne Midgette" and "mother" are substituted by "she". In general, pronouns refer back to nouns in the discourse. In that sense, they are phase 2 operators. But the pronouns "I" and "you" do not anaphorically refer back to a noun in the discourse: they are linked to the act of communication between the speaker and the co-speaker. In general, personal pronouns are phase 2 operators. This analysis establishes that the linguistic units known as grammatical references can be organized on the phase systems. Their metalinguistic value disclosed are differential properties on which the teacher can focus to teach them. The reference to the extra-linguistic to explain them can cause confusion. Therefore, it is necessary to disclose their metalinguistic values. The following table summarizes the grammatical forms studied and their metalinguistic values:

Table 2. Operators and Metalinguistic Values

Operators	Metalinguistic functioning
A; this	Phase 1 operator: used to introduce new information
The; that; Personal pronouns[6]	Phase 2 operator: used to recall information

Conclusion

Cohesive writing can be a conundrum for English learners if they are not presented with the appropriate linguistic tools that help them understand the internal working of language. Our approach to grammatical cohesion is metalinguistic oriented. This study shows that metaoperational grammar theory and corpus linguistics can be used as a bedrock to help students produce cohesive writing in order to succeed in conveying their own ideas. While corpus linguistics offers millions of utterances to teachers, metaoperational grammar gives them the necessary analytic tool to look into natural language to explain its functioning. Unlike the traditional way to collect data, this method is much faster, reliable and representative. We can reach millions of utterances to process. It is therefore important for the university institutions of Côte d'Ivoire to take advantage of this fast-growing methodology in language teaching.

---

[6] The pronoun "it" deserves a special treatment because in some cases, it has no clearly identified reference. A case in point is: *it is cold to day. It rains* (J.-R. Lapaire and W. Rotgé, 1991, p.215).

Bibliography

Adamczewski, H. & Gabilan, J-P. (1992). Les Clés de la grammaire anglaise. Paris : Armand Colin.

Adamczewski, H. and Claude, D. (1982). Grammaire linguistique de l'anglais. Paris : Armand Colin.

Adamczewski, H. (1976). Be+-ing dans la Grammaire de l'Anglais Contemporain, Thèse d'état sous la dir. d'Antoine Culioli, Université de Paris VII, [version publiée, Paris, Honoré Champion, 1978].

Ahmed, A. R. (2008).Reference as a Cohesive Device", *ADAB AL-RAFIDAYN*, (52)43-68

Cook, G. (1989). Discourse. Oxford: Oxford University Press.

Gabilan, J-P. (2011). L'imparfait français et ses traductions en anglais : approche méta-opérationnelle, coll. *Langages*, Université de Savoie, Chambéry

Gries, S. T. (2009). What is Corpus Linguistics, *Language and Linguistics Compass*, (3), 1225-1241

Halliday, M. A. K. & Hasan, R. (1976). Cohesion in English. London: Longman.

Hassani, G. (2011). A Corpus-Based Evaluation Approach to Translation Improvement, Meta, 56(2): 351–373. [En ligne], consultable sur DOI : https://doi.org/10.7202/1006181ar

Hinkel, E. (2001). Matters of cohesion in L2 academic texts. Applied Language Learning 12(2) : 111–132.

Hoyoon, E. & Byoungman, J. (2019). Reference and Substitution as Cohesion Devices in EFL Writing, *English Language and Literature Education*, (15)4:23-36. [En ligne], consultable sur URL : https://koreascience.kr/article/JAKO200917357306504.pdf

Johnson, M. (2017). Improving Cohesion in L2 Writing: A Three-Strand Approach to Building Lexical Cohesion. [En ligne], consultable sur URL :https://files.eric.ed.gov/fulltext/EJ1166137.pdf

Kahlaoui, M. H. (2019). Metaoperational Linguistics: issues of translatability and Visibility, *Arab World English Journal*, 10 (2) 3-15. [En ligne], consultable sur DOI: https://dx.doi.org/10.24093/awej/vol10no2.1

Kpli, Y. K. J-F. (1992). A Grammatical Exception as a Construct, in *Cahiers Ivoiriens de Recherche Linguistique (CIRL)*, 26, ILA.

Lachaux, F. (2006). Imparfait ou imparfaits ? Mise en perspective d'un marqueur complexe : approche méta-opérationnelle, étayée par quelques éléments de contrastivité, *Actes du colloque de juin 2006, Crelingua*, EMA éditeur.

Mahlberg, M. (2009). Lexical cohesion: Corpus linguistic theory and its application in English language teaching." In *Lexical cohesion and corpus linguistics*, ed. J. Flowerdew and M. Mahlberg. Philadelphia: John Benjamins, 103–122

Meyer, C. (2004). English Corpus Linguistics: An Introduction. Cambridge: Cambridge University Press.

McEnery, T. & al. (2006). Corpus-Based Language Studies. An Advanced Resource Book (Routledge Applied Linguistics). London, New York: Routledge.

McEnery, T. & Hardie, A. (2012). Corpus Linguistics Method, Theory and Practice.

Cambridge: Cambridge University Press

Nordquist, R. (2019). Definition and Examples of Corpus Linguistics. [En ligne], consultable sur URL :https://www.thoughtco.com/what-is-corpus-linguistics-1689936

Pantaleon, A. M. (2018). A corpus-based analysis of "for example" and "for instance", *The Asian ESP Journal*, (14)7.2:151-175

Seleskovitch, D. & Lederer, M. (2014). Interpréter pour traduire. Paris : Les Belles Lettres.

Swales, J. & Feak, C. (2004). Academic writing for graduate students: Essential tasks and skills (2nd ed.). Ann Arbor, MI: The University of Michigan Press.