

A corpus linguistics application in the analysis of textbooks as national teaching instruments of English as a Second Language in Chile

Aplicación de la lingüística de corpus en el análisis de libros de texto como instrumentos nacionales de enseñanza del Inglés como segunda lengua en Chile

Volumen 18, Número 1

Enero-Abril

pp. 1-19

Este número se publica el 1° de enero de 2018

DOI: <https://doi.org/10.15517/aie.v18i1.31807>

Rodrigo Arellano A.

Revista indizada en [REDALYC](#), [SCIELO](#)

Revista distribuida en las bases de datos:

[LATINDEX](#), [DOAJ](#), [REDIB](#), [IRESIE](#), [CLASE](#), [DIALNET](#), [SHERPA/ROMEO](#),
[QUALIS-CAPES](#), [MIAR](#)

Revista registrada en los directorios:

[ULRICH'S](#), [REDIE](#), [RINACE](#), [OEI](#), [MAESTROTECA](#), [PREAL](#), [CLACSO](#)

A corpus linguistics application in the analysis of textbooks as national teaching instruments of English as a Second Language in Chile

Aplicación de la lingüística de corpus en el análisis de libros de texto como instrumentos nacionales de enseñanza del Inglés como segunda lengua en Chile

Rodrigo Arellano A.¹

Abstract: The Ministry of Education (MINEDUC) bases the instruction of English in Chile in the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFRL), however its guidelines have not been systematically considered in the practice of creating instructional material in the EFL context of this country. To analyze this issue, this article presents the comparison of the vocabulary about Teaching English to Speakers of Other Languages (TESOL) used in the CEFRL and the English textbooks used in Chile regarding the categories of curriculum design, language teaching methodology and language assessment. The methodology was mixed, not experimental, cross-sectional and descriptive, including the analysis of the CEFRL document and the 8 English textbooks used in Chile in the public sector. Firstly, the data was compared quantitatively using the software Nvivo9 and later some key words (learners / students - teachers) were analyzed quantitatively in terms of frequency and qualitatively using collocations. Results suggest that similar vocabulary is used in both, the CEFRL and Chilean textbooks in terms of "language teaching methodology" (36% and 33% respectively), but it varies highly in the category of "curriculum design" (22% and 48% respectively) and mostly in "language assessment" (42% and 19% respectively) while showing different frequencies in the key words and their associated verbs between "students" (1,33%) and "teacher" (0,18%). These results show a constructivist approach in both, but with a minor behaviorist aspect especially in textbooks, indicating relevant differences in the emphasis and coherence between different areas of the instruction.

Key words: teaching guides, CEFRL, teachers, students.

Resumen: El Ministerio de Educación (MINEDUC) fundamenta la instrucción de la enseñanza del Inglés, en Chile, en el Marco Común Europeo de Referencia para las Lenguas: Aprendizaje, Enseñanza, Evaluación (MCER), sin embargo, sus directrices no han sido sistemáticamente consideradas en la práctica de crear material instruccional en Inglés en este país. Para analizar esta problemática, este artículo presenta la comparación del vocabulario sobre la enseñanza del Inglés para aprendices de otras lenguas (TESOL siglas en Inglés), usado en el MCER, y los libros de texto empleados en relación con las categorías de diseño curricular, metodología de la enseñanza del idioma y evaluación de la competencia lingüística. La metodología fue mixta, no experimental, transversal y descriptiva, en la cual se incluyó el análisis del MCER y de los 8 libros de Inglés utilizados en el sector público durante el año 2011. Primero, los datos fueron comparados cuantitativamente mediante el software Nvivo9 y posteriormente fueron analizadas de manera cuantitativa (frecuencias) y cualitativa (colocaciones) algunas palabras clave (aprendiz / estudiante - profesor/a). Los resultados sugieren que se usa vocabulario similar en ambos corpus (MCER y libros de texto) en términos de metodología (36% y 33% respectivamente), pero varía altamente en currículo (22% y 48% respectivamente) y mayormente en evaluación (42% y 19% respectivamente) ello muestra diferentes frecuencias en las palabras claves y los verbos asociados entre "aprendiz-estudiante" (1,33%) y "profesor/a" (0, 18%). Estos resultados muestran un acercamiento constructivista, aunque con aspectos conductistas, especialmente en los libros de texto, lo cual evidencia diferencias relevantes en el énfasis, y coherencia entre las áreas de conocimiento.

Palabras clave: manuales de enseñanza, MCER, docente, estudiante.

¹ He is currently a lecturer in the Department of Languages, Literature and Communication at La Frontera University (Temuco, Chile). Master's degree in Applied Linguistics and Master's degree in Educational Studies, University of New South Wales, Sydney, Australia.

E-mail address: rodrigo.arellano@ufrontera.cl

Artículo recibido: 16 de mayo, 2017

Enviado a corrección: 28 de setiembre, 2017

Aprobado: 13 de noviembre, 2017

1. Introduction²

The Common European Framework of References for Languages (CEFRL) provides a set of guidelines to understand how languages can be taught and learnt (Council for Cultural Co-operation, 2001). This framework is used not only in the European Union, but also by other countries such as Chile, which has a state policy agenda to improve the quality of English teaching in the country.

In particular, the Chilean Ministry of Education created the English Open Doors program to reinforce English teaching to "provide the students with tools to access the technical, commercial, cultural and academic world" (Ministry of Education, 2002, p. 41). Furthermore, one of the backbones of the English Open Doors program is the provision of free English textbooks (MINEDUC, 2009), based on the guidelines proposed by the CEFRL, which are highly influenced by the Communicative Language Teaching Approach.

In this context, this article will report the extent in which the English textbooks in Chile reflect the European philosophy of teaching languages according to the teachers' guidelines of the three-part pedagogical process: planning, implementation and assessment (Ander - Egg, 1996). Firstly, the general objective of it is to compare the vocabulary proposed by the Council for Cultural Co-operation with its realization in English textbooks regarding Curriculum Design, Language Teaching Methodology and Language Assessment. And specifically, the research questions are a) to what extent is vocabulary about curriculum design in the CEFRL reflected in the design of English textbooks in Chile? b) how are the concepts from the CEFRL, reflected in Chilean English textbooks regarding TESOL methodology? c) to what extent language assessment instruments are consistent with the way Chilean textbooks apply the theory provided by the Council for Cultural Co-operation? Finally, it is important to notice that not only percentages were used to answer these questions, but also collocations of key words in order to understand the data in a better way.

2. Theoretical framework

2.1 Philosophy behind CEFRL

The CEFRL considers language as a powerful social construct. Wittgenstein (1972) explains this position with the notion of 'language - game', as a way of understanding reality

² This work was supported by the Scholarship - "Becas Chile" from The National Commission for Science and Technological Research (CONICYT - Chile) as one of the requirements to obtain the Master's degree in Applied Linguistics (2012) at the University of New South Wales (Sydney - Australia).

through communication. He interprets language as a complex web of relations between verbal and non - verbal factors that are governed by rules and involve diverse models of reality perceived by different interlocutors. This is not only relevant to discover the philosophy behind the CEFRL, but also "to provide a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc." (CEFR, 2001, p.1).

One important notion behind the CEFRL comes from the concept of communicative competence, which is divided into linguistic, sociolinguistic and pragmatic capabilities. According to Fulcher and Davidson (2007) 'linguistic competence' comes from the Chomsky's terminology and relates to everything speakers know about language. Then, 'sociolinguistic competence' is based on the social rules set in a particular context of speakers following Hymes' argument that a Chomskyan approach is true, but insufficient. Finally, 'pragmatic competence' refers to strategies in the form of speech acts, or things we do with words. These three levels of communicative competence have influenced the way language teaching is conceived not only in units of linguistic analysis, but also in context.

Regarding a user's language competence, the CEFRL provides a framework divided into three categories to express what is involved in a learner's level of proficiency:

- a. General competence: This type of competence refers to the person's knowledge of the world, how this knowledge is used socioculturally and the importance of intercultural awareness.
- b. Skills and know - how: This kind of competence deals with the practical use of the languages learnt and how this learning is influenced by the convergence of different cultures.
- c. Existential: This is connected with personal factors that shape the learner's linguistic performance. These include attitudes, motivation, values, beliefs, cognitive styles and personality factors.

Additionally, tasks and their role in language teaching become relevant due to its recurrence in language instructional settings, especially textbooks. According to Richards (2001) "a task is an activity or goal that is carried out using language" (p. 161) considering competences, conditions - constraints and strategies. Regarding task difficulty, language teachers should also bear in mind that cognitive and affective factors can influence the task performance and that both comprehensive and productive tasks are shaped by the context and the genre being used. This is particularly relevant since pedagogic materials are often

regulated not only in terms of linguistic complexity but also in task difficulty, having to represent authentic language samples and contexts of use most of the time (Brown, 2005).

Another aspect of the CEFRL is the importance of learning and teaching, since this document was the result of years of research to orient the way educational practices were done in Europe. From this approach, the CEFRL assigns importance to students as central participants of the learning process, taking into consideration factors such as students' needs, personal abilities and multilingual / multicultural competences. From this view, the framework explores different approaches to reflect on how languages are learnt and acquired, although the Council of Europe itself does not choose any approach in particular for a linguistic pedagogical practice. Or in Harmer's words (2007) the aim is to decide "one approach for context-sensitive teachers to try to create a bridge between their methodological beliefs and the students' preferences" (p. 77). Finally, it emphasizes the usage of the CEFRL guidelines for diverse members of the language teaching community.

However, one of the most remarkable and known contributions from the CEFRL in the European context is the agreement among many speech communities of common reference levels for describing language proficiency (Jiménez, 2004). Hence and according to the CEFRL, language proficiency is divided into three levels, outlined in the table 1:

Table 1. Classification of the levels of language proficiency according to the document of the Common European Framework of Reference for Languages: Learning, Teaching, Assessment

User	Proficiency level
A.Basic	Breakthrough (A1) Waystage (A2)
A.Independent	Threshold (B1) Vantage (B2)
B.Proficient	Effective operational proficiency (C1) Mastery (C2)

Source: Adapted from Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Strasbourg: Press Syndicate of the University of Cambridge.

This scale also provides a detailed description of what the learner is able and not able to do with the language (can - do statements) (O'Dowd, 2010). This is done so within communicative strategies in language understanding (reading / listening) and production (writing / speaking), in such a way that they become coherent with standardized descriptors of international language testing systems (Hughes, 2003). Can - do statements help teachers to identify the areas students master and the ones that they do not and they are divided into communicative themes, tasks / purposes, activities, strategies, processes and texts

representing different domains of language, with categories that are strictly related to planning, implementing and assessing language performance in the European context.

2.2 Planning, teaching and assessing as principles within the pedagogical practice

Firstly, curriculum design refers to planning goals, learning outcomes and gradation / sequencing of content that are important in organizing the learning process catering to achieve educational goals established within an organized plan (Richards, 2001). In the same way, contents are sequenced following complexity and difficulty principles in terms of language skills and morphosyntactic elements (O´ Dowd, 2010).

Secondly, language teaching methodology is commonly referred to the teachers' pedagogical practice or classroom practice (Nunan, 2004) and it deals with the range of the educators' activities: to deliver language content, to control classroom management and to design pedagogical material (Ur, 2000) whilst considering individual differences and social factors that influence students' needs (Graves, 2001).

Finally, assessment refers to types of evaluation by purpose, focusing on achievement and proficiency assessment and also including rating, judgement and criteria from the specification of tests, the criteria for the attainment of a learner objective and the learner's levels of proficiency (Alderson, Clapham and Wall, 1995; Bachman and Palmer 1996; Brown, 2005; Fulcher and Davidson, 2007). Traditionally, it has been stated that tests have to measure what they are supposed to measure (validity), they must provide similar testing results when used in different occasions (reliability) and they have to be efficient in terms of time, cost and effort (feasibility) (Hughes, 2003).

2.3 English teaching in Chile

During the 90's all Latin-American countries commenced educational reforms with the purpose of increasing the efficacy and quality of their educational processes in which educational approaches were rethought and redesigned (Barahona, 2016).

This educational reform was implemented in Chilean schools changing the paradigm from behaviourism to constructivism. According to Kaufman (2004) "the emergence of this paradigm has coincided with a shift in pedagogy away from teacher-centred information transmission models toward knowledge centred approaches that focus on cognitive and social processes in learning" (p. 303). Precisely and in relation to the teaching of the English

language in Chile, the majority of the teachers used the grammar-translation method or even audio-lingualism techniques in their pedagogical practices before moving into more contextual approaches (McKay, 2003). Hence, when constructivism came onto the scene, this structuralist view was changed into the communicative language teaching approach (Barton, 1994) and it is still in process (Barahona, 2016).

Then in this new approach, the school was supposed to prepare learners for real life situations, especially in the EFL area where there is a need to master a common language of communication (Richards, 2001), as there is more and more agreement on the purposefulness of education and schooling as a way of socializing within a culture (Barber and Mourshed, 2007). Thus, English started to play a special role in this instructional process; since it would help students to go beyond frontiers in the world especially considering that Latin America has been "profoundly affected by the modernity - modernization and the globalization process derived from these changes" (Pérez, 2009, p.30).

In particular and among the initiatives of the Ministry of Education of Chile to improve the teaching of the English language in the country, there is one that is particularly relevant for English Teaching as a foreign language: "The English Open Doors Program" (Barahona, 2016). And more specifically, one of the most important policies of this department was the provision of free English textbooks. This new practice was relevant as according to Byrd (2001), "most teachers depend on textbooks - often as a required tool - because they provide content and activities that shape what happens in the classroom" (p. 415). Therefore, textbooks are important as they help the teacher to have a progression of contents with clear learning outcomes coherent with the state of the field and the guidelines provided by national governments, including the ones provided by the CEFRL. In the Chilean experience, Bitar (2002), former Minister of Education, stated that "English textbooks have been delivered to all children currently in the fifth and sixth grades in municipal and private subsidized schools, along with a tape for each course" (p. 2). This might be a simple strategy for some countries, but it was a real revolution in South America since every single child received material in English for the very first time.

3. Methodology

3.1 Research approach

Word queries were used to find the frequency of TESOL - related vocabulary in a quantitative way (automatic analysis) and qualitative analysis in the form of collocations was

also utilized (manually). First and in terms of the former approach, Lee (2007) has explained that software can increase the trustworthiness of research process and therefore they represent an objective strategy to manage data so that the foundation of the study can be solid through numerical information from the very start. Additionally, some data was analysed qualitatively because according to Denzin and Lincoln (2005) qualitative research "makes the world visible in a different way." (p. 4) Hence, an advantage of a quantitative approach with a minor qualitative component is that it is not only interested in stating frequencies or text patterns through numerical information or statistics. But most importantly, the interest is on what the results represent, whilst providing useful insights about textbook authors' philosophies behind their construction in terms of the interpretation that the information can provide.

3.1 Design

A corpus-based study was carried out to compare the vocabulary used in the CEFRL and that used in Chilean textbooks to discover whether the word choices were replicated or not in these two sets of data through a descriptive type of design, non-experimental and with a cross-sectional nature. Therefore, the aim of the analysis is to show what the data contains in order to unmask the characteristics of the phenomena in order to illustrate how the reality is.

3.2 Materials

Chilean textbooks designed and used by the Ministry of Education in 2011 and the Common Framework of Reference for Languages (2001) were compared and analysed using software designed to analyze corpora. It is important to notice that the students' textbooks were analyzed in light of the teachers' guides since this version of the book offers guidelines for the instructional process with longer and richer texts full of linguistic and pedagogical / technical vocabulary. Additionally, the books were analyzed in total representing the 8 years Chilean students attend English courses, that is, primary school (4 years) and high school (4 years).

3.3 Instrument

The analysis software Nvivo9 was used to find recurrent text patterns in both the textbooks used by students and the pre - selected chapters from the Common Framework of

Reference for Languages. According to Leech and Onwuegbuzie (2011) this software is particularly designed to manage non-numerical and unstructured data building tools to identify trends, highlight key points, classify information into themes and examine the relationship among diverse pieces of data.

3.4 Procedure

First of all, eight Chilean textbooks for teaching English and the Common Framework of Reference for Languages were analyzed quantitatively in terms of word frequency. They were stored in the Nvivo9 software so that repeated key words were found. These key words were categorized into three broad categories: curriculum design, language methodology and language assessment, in order to compare the word choices in the CEFRL and textbooks. After this analysis was done separately (textbooks and chapters in different analyses), the different categories and their corresponding word frequency were compared. Finally, collocation analysis was used to explore key words that were relevant for the TESOL profession (learner / student - teacher) in light of the three broad categories. This is relevant since "this combination of qualitative familiarity with the data and quantitative analysis allows the researcher to make interesting findings with immediate pedagogical implications" (Lee, 2007, p.90) to be guided by the research questions to discover the consistency between curriculum, methodological and assessment concepts between the CEFRL document and the English textbooks in Chile.

Additionally, it is important to clarify that corpus analysis works in an opposite way compared to traditional scientific methods. Nunan and Bailey (2009) explain that "inductive reasoning begins with data and ends up with a theory" (p. 421) as opposed to deductive theory that begins with a theory and then confirms it or refuses it. Therefore, from a methodological point of view, data analysis is done via a bottom - up approach. For this, both sets of documents were compared to find reasons for matches or mismatches regarding similarities and differences in meaningful words in light of the Chilean experience of English teaching.

4. Results and discussion

4.1. Categories and subcategories found in the data

First, and after running the software the vocabulary was classified into three categories with ten subcategories in total. The general results are shown in the table 2:

Table 2. Classification of the categories and subcategories of the TESOL - related vocabulary in the CEFRL and the English textbooks used in the public sector in Chile in 2011

Curriculum Design	Language Teaching	Methodology	Language Assessment
Syllabus design	Text delivery	Test types and items	
Course planning	Classroom management	Evaluative vocabulary	
Tasks	Qualifying skills		
Materials and resources	Feedback		

Source: Researcher's own design (2013).

As it can be seen above, three main categories were found regarding the vocabulary in the CEFRL and English textbooks. Then, subcategories were created to order the data and finally these subcategories were edited manually. The limitation of this process was based on the impossibility to categorize some concepts that may belong to any of the categories expressed in the research questions (Härmälä, 2010). To solve this problem, language choices difficult to analyse were categorized according to their context (collocations) and in some cases, these concepts were assigned to more than one category (Yu, Jannasch-Pennel and Di Gangi, 2011).

Firstly, a frequency query was run and then TESOL vocabulary was categorized in these three initial categories. The results of this analysis are summarized in table 3 below:

Table 3. Comparison of the percentages for the use of TESOL-related vocabulary in the three categories of the study

Set of data (document)	Curriculum design	Language teaching methodology	Language assessment
CEFRL	22%	36%	42%
Textbooks	48%	33%	19%

Source: Researcher's own design (2013).

4.2 New categories emerging from the original data

In qualitative methodology, the researcher is conceived as part of the investigation process, thus the criterion to place concepts in different categories is based on the researcher's knowledge of the data (Mella, 2003). However, the original main categories were extended to five to include the categories 'linguistic technical vocabulary' and 'educational vocabulary' with lexicon placed into one of the original categories. Considering the fact that "one of the main concerns regarding corpus-based analysis is their lack of contextual features for interpretation of the corpus data" (Flowerdew, 2008, p. 115) collocation (manual) analysis

was used with the 15 first collocations in order to move those concepts into one of the three original categories.

4.3 Concepts with the highest frequency in the data

Additionally, the word query shows the most repeated concepts across categories as it can be seen in table 4 below:

Table 4. Summary of the TESOL-related tenets with the number of reference for key terms and their percentages across categories

Category	Word	Number of references (times)	Coverage in the document (percentage)
Curriculum design	Task (CEFRL)	354	0.49%
	Units (textbook)	1381	0.43%
Language teaching methodology	Writing (CEFRL)	238	0.33%
	Activity (textbook)	277	0.38%
Language assessment	Descriptors (CEFRL)	291	0.40%
	Correct (textbook)	829	0.26%

Source: Researcher's own design (2013).

Finally, the three research questions of this study were answered one by one:

- a) To what extent is the theory from the Common European Framework of References for Languages reflected in the design of unit plans in English textbooks in Chile?

In general terms, the vocabulary contained in the CEFRL has a minor representation in the textbooks used in Chile, in which the frequency of key concepts is much higher, although it does include a minor percentage of TESOL-related vocabulary which is inexistent in the Council for Cultural Co-operation's proposal. In other words, CEFRL only contains 22% of TESOL - related lexis whereas curriculum vocabulary is almost half of the word choices used in textbooks (48%).

If the vocabulary from curriculum design is compared across subcategories, there is a high correlation of technical concepts, differing only in 6% of the data. However, there are some curriculum terminologies that belong to only one of the sets of data. In this case, the CEFRL presents concepts that have been placed in different categories based on collocation analysis. An interesting finding is that the most repeated concept in the CEFRL has been the word 'task' mentioned 354 times with a 0.49% of the total weighted percentage of the data. On the contrary, the concept 'units' is the word with the highest frequency in the textbooks

being mentioned 1381 times with a 0.43% of weighted percentage. Finally, the results by subcategories in 'curriculum design' are outlined below in table 5:

Table 5. Classification of the subcategories created to group the curriculum-related vocabulary and its distribution in the data

Set of data (document)	Curriculum design	
	Syllabus design	Course planning
CEFRL	41%	59%
Textbooks	47%	53%

Source: Researcher's own design (2013).

- b) How are the principles from the CEFRL taken into account in textbooks to implement English lessons in Chile?

Results suggest the vocabulary used in the CEFRL is replicated in the textbooks used in Chile with minor differences between the two sets of data (36% in the CEFRL and 33% in textbooks). However, if subcategories are analyzed, it is possible to find differences in the words used as well as their frequencies. For example, the category 'tasks' in the CEFRL shows a more significant difference in comparison to the vocabulary in the textbooks used in Chile. The words with the highest frequency are 'activity' in the textbooks (mentioned 277 times and representing a 0.38% of the weighted percentage) and 'writing' in the CEFRL (mentioned 238 times with a 0.33% of the weighted percentage) which suggests a strong emphasis from the CEFRL on activities related to written production. On the contrary, textbooks seem to emphasize receptive skills rather than productive skills. This can be seen in the query in which the two concepts classified under the category 'language teaching methodology' represent the highest frequency. These two concepts are 'reading' (mentioned 2.500 times representing 0.78% of the text) and 'listening' (mentioned 1996 times representing 0.62% of the text) suggesting textbooks in Chile provide more emphasis on these skills than writing. This finding is especially relevant since speaking in both sets of data represents only 0.05% of the texts. Finally, the results by subcategories in 'language teaching methodology' are outlined below in table 6:

Table 6. Classification of the subcategories created to group the methodology-related vocabulary and its distribution in the data

Set of data (document)	Language teaching methodology			
	Materials and resources	Text and mode delivery	Classroom management	Tasks
CEFRL	27%	9%	17%	47%
Textbooks	26%	9%	14%	51%

Source: Researcher's own design (2013).

- c) In what manner do language assessment instruments apply the theory provided by the Council for Cultural Co-operation?

The greatest weakness of the textbooks, at least from a corpus - based perspective, is the treatment of assessment, since it shows the lowest correlation among the three categories (42% in the CEFRL and 19% in textbooks).

Similarly, among the three compared categories, vocabulary related to assessment represents the highest difference in subcategories as well, especially in 'evaluative vocabulary' and 'test types and items'. Additionally, evaluative vocabulary represents 39% of the language choices for assessment in CEFRL compared to only 19% in the textbooks.

The CEFRL contains more specific vocabulary to judge a learner's competence (e.g. strengths, weaknesses, badly, etc.) while textbooks commonly use 'classroom language' (e.g. good, bad, correct, incorrect, etc.). The CEFRL also uses more synonymy than the textbooks which repeat the same adjectives to assess language competence. Finally, the words with the highest frequency in the CEFRL and textbooks are 'descriptors' (mentioned 291 times with a 0.40 of weighted percentage) and 'correct' (mentioned 829 times representing the 0.26% of the text) respectively. Finally, the results by subcategories in 'language assessment' are outlined below in table 7:

Table 7. Classification of the subcategories created to group the assessment-related vocabulary and its distribution in the data

Set of data (document)	Language assessment			
	Text types and items	Qualifying skills	Feedback	Evaluative vocabulary
CEFRL	17%	22%	22%	39%
Textbooks	31%	25%	25%	19%

Source: Researcher's own design (2013).

One particular set of words that are of interest in the analysis are 'learner/student' in contrast with the word 'teacher' in order to find the emphasis both sets of documents are providing on who the protagonists in the learning - teaching process are in relation to curriculum design, language teaching methodology and language assessment³. The frequency of these concepts is outlined below in table 8:

Table 8. Summary of the number of references and coverage of the key words from the data

Key words	Set of data (document)	Number of reference (times)	Coverage in the document (percentages)
Learner	CEFR	560	0.32%
	Textbooks	227	0.03%
Teacher	CEFR	132	0.07%
	Textbooks	898	0.11%
Student	CEFR	30	0.02%
	Textbooks	5241	0.66%

Source: Researcher's own design (2013).

Firstly, the most salient feature when analyzing this table is the high frequency of the concepts 'learner' and 'student' when compared to the word 'teacher'. This is a clue to assume that both sets of documents are based on a constructivist approach in which the learner is the real protagonist of the instructional process and the teacher is conceived as a facilitator or mediator (Kalaja and Ferreira, 2003).

Nonetheless, the word query in the textbooks has a higher frequency on the side of the learner due to the high frequency of the word 'student'. Since textbooks are based on the application of a curriculum within an educational institution, it is logical to think that the language options used to describe the learner are contextualized within that institution. However, Pérez (2009) states that institutions tend to produce and reproduce culture to design curriculums based on social hierarchies, so that social order can be maintained and replicated. Therefore, the word 'learner' is replaced by a more concrete synonym (student) that denotes learning from a behaviorist position. As opposed to the textbooks, the CEFR uses this concept in an opposite fashion. Since the Council for Cultural Co-operation encompasses language learning not only in a school, but in other academic scenarios (including non - academic such as the familiar background), the word learner is used including any individual wanting to learn and assess his or her language proficiency, whether

³ This analysis with collocation patterns is useful because they can be investigated in the context of semantic prosody to find negative or positive evaluative meanings (Flowerdew, 2005).

independently or in school contexts. Although this notion is not exclusive, it is more in line with an autonomous concept of learning embraced by constructivist theories in which educational institutions are a medium and not the end of learning. This coincides with a move towards the learner - centered curriculum, in fact, the word 'learner' is the fourth most repeated word after 'can't', 'using' and 'language' representing 0.79% of weighted percentage in the CEFRL.

Another relevant aspect to be considered is the usage of the word 'teacher' in both sets of documents. The frequency of the query is low in the CEFRL and in textbooks, but no other possible synonyms such as 'instructor' or 'facilitator' are used. Even when both sets of data refer to tasks in class, methodology or assessment, the teacher is always conceived as a human resource for the learner / student to carry out tasks or to achieve objectives. However, according to Orafi and Borg (2009) curriculum innovations should not only consider teachers' practices and beliefs, but also should allow the teacher to suggest curriculum innovations.

Regarding categories in which concepts were classified, the learner is considered in the planning and implementation stages of the instructional process, while the teacher is commonly placed functionally to service student's tasks while creating the conditions in which learning can take place (Scrivener, 2005). Again, this falls in line with constructivist positions of language learning in which the student is the centre of the pedagogical actions. These relations are summarized in the table 9:

Table 9. Relationship of the concepts of learner - student and teacher across different subcategories

Concept	CEFRL (sub-categories)	Textbook (sub-categories)
Learner	Feedback	Syllabus design
Student	Feedback	Text delivery, Evaluative vocabulary
Teacher	Syllabus design, Classroom management, Evaluative vocabulary	Materials and resources, Feedback

Source: Researcher's own design (2013).

As it is outlined above, 'learners / student' are placed in language assessment categories in the CEFRL and in curriculum design and language teaching methodology in textbooks. This seems logical as the CEFRL defines descriptors to place students in different levels since feedback is an essential component of testing and evaluation. On the contrary, planning and implementation of the pedagogical process are emphasized more in textbooks since pedagogic aids for students are more concrete in terms of instructional objectives and evaluative vocabulary and they are only used to judge good or bad performances in

'classroom language'. Again, a very interesting aspect to be considered is that of 'teachers' being perceived as a human resource when planning, implementing and assessing the language learner. Despite this, they are not considered in the planning stage in textbooks. This might suggest the functional idea the Ministry of Education has of teachers since their role would not be to intervene on curriculum design, but only in implementation and assessment of language performance (Canagarajah, 1993).

As a final point, there are two aspects that are notable in the analysis since they demonstrate a contrast between constructivist and behavioural theories of language learning. The CEFRL places the teacher in the category of 'classroom management' as the instructor is normally in charge of controlling disruptive behaviour (Ur, 1999). However, no references are done into student's self - discipline (Guilloteaux and Dörnyei, 2008), but only regarding discipline regulated by others (Deci and Ryan, 2000) without referring to strategies that make the learner an autonomous protagonist in the classroom (Garret and Shortall, 2002). Rather, collocation analysis shows that the teacher is still seen as the center of the action in the class, which is unquestionably a behaviourist position towards pedagogy (Lantolf and Poehner, 2004).

However, textbooks take it a step further when referring to the category 'text delivery' in which the students are seen as the protagonists to produce, reproduce or create texts. Therefore, there might be a conflict regarding who the actual protagonist of the classroom is as "resistance might come from cultural notions of the respective roles and responsibilities of teachers and learners which have fostered certain habitual styles of learning" (Hedge, 2000, p. 100), especially considering the fact these texts emphasize listening and reading skills rather than writing and speaking. In other words, these two ideas reflect a change of paradigm in which students - learners and teachers must behave differently despite the fact their roles are expected to occur in a particular way (Cook, 2001).

To sum up, frequency of words suggests the learner / student is the centre of the action, but collocation analysis provides useful insight to notice that although there is a constructivist position behind the CEFRL and textbooks, behaviorism ideas are still present in proposals for classroom scenarios.

5. Conclusions

The findings of this study explore the replication of the vocabulary used in the CEFRL with the language choices used in English textbooks in Chile, and particularly with some key

words in the texts used by teachers. Results suggest a correlation in the type of TESOL vocabulary used in the data, however word frequency varies significantly, especially in the areas of curriculum design and language assessment, but it is relatively similar in the category of language teaching methodology. Results suggest that similar TESOL - related vocabulary is used in the CEFRL and Chilean textbooks to teach English, however the frequency of these concepts varies highly regarding curriculum design and language assessment concepts the frequency of words is similar regarding language teaching methodology. These results can be explained as textbooks value more the tasks to be done in class (language teaching methodology), but not the planning (curriculum design) and the evaluation process (language assessment).

These different frequencies reflect a diverse degree in the philosophies behind TESOL principles (Council for Cultural Co-operation, 2001). And therefore, an emphasis in some areas in the CEFRL, especially curriculum design and language assessment, whose treatment seems weaker in the textbooks, might be affected when teachers use this material considering the influence textbooks have in the teachers' pedagogical practices (Byrd, 2001). Besides that and in relation to the key words from the analysis, a constructivist conception of teaching is shown in the texts (Kaufman, 2004), despite the fact behaviorist aspects of the pedagogical process are still found in both learners and teachers and in particular, from the perspective of the textbooks used in Chile to teach English as a foreign language.

Despite the effort undertaken, limitations can be found in this study. First of all, the data is limited to a small amount so the study is restricted to a comparison between only two sets of documents. Additionally, corpus analysis is often decontextualized so the resulting collocation analysis is useful although it is time consuming and can be interpreted rather subjectively at times. Therefore, corpus analysis works from a bottom - up approach not considering genre or discourse level.

Notwithstanding, this work can be the foundation for further research in text analysis with the purpose of understanding the beliefs and ideologies of the producers of pedagogical texts. This can be done by integrating more texts to build a more representative corpus or to compare this corpus to other texts that show similarities in terms of their teaching approaches. Also, the exploitation of corpora through software can be used to analyze lesson plans and the emphasis given to the key words analyzed in this study and even classroom talk can be transcribed to have automatic analysis of the participants within the L2 classroom.

6. References

- Alderson, J. Charles, Clapham, Caroline and Wall, Dianne. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Ander-Egg, Ezequiel. (1996). *La planificación educativa: conceptos, métodos, estrategias y técnicas para educadores*. Buenos Aires, Argentina: Editorial Magisterio del Río de la Plata.
- Bachman, Lyle, F., and Palmer, Adrian, S. (1996). *Language testing in practice. Designing and developing useful language tests*. Oxford: Oxford University Press.
- Barahona, Malba. (2016). Challenges and Accomplishments of ELT at Primary Level in Chile: Towards the Aspiration of Becoming a Bilingual Country. *Educational Policy Analysis Archives*, 24(82), 1-29.
- Barber, Michael and Mourshed, Mona. (2007). *How the World's Best-Performing School Systems Come Out On Top*. s.l.: McKinsey and Company.
- Barton, David. (1994). *Literacy: An Introduction to the Ecology of Written Language*. Oxford: Blackwell.
- Brown, Annie. (2005). Self-assessment of writing in independent language programs: The value of annotated samples. *Assessing Writing*, 10(3), 174-191.
- Byrd, Patricia. (2001). Textbooks: Evaluation for selection and analysis for implementation (4th ed.). In Marianne Celce-Murcia, M. Donna Brinton and Ann Marguerite Snow (Eds.), *Teaching English as a second or foreign language* (pp. 415-427). Boston: Heinle, Cengage Learning.
- Canagarajah, A. Suresh. (1993). American textbooks and Tamil students: Discerning ideological tensions in the ESL classroom. *Language, Culture and Curriculum*, 6(2), 143-156.
- Cook, Vivian. (2001). *Second Language Learning and Language Teaching*. London: Edward Arnold.
- Council for Cultural Co-operation. (2001). *Common European framework of reference for languages: Learning, teaching, assessment: Education Committee. Modern Languages Division*. Strasborg: Press Syndicate of the University of Cambridge.
- Deci, Edward and Ryan, Richard. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, 55(1), 68-78.
- Denzin, Norman and Lincoln, Yvonna. (2005). *The Sage handbook of qualitative research*. California: SAGE Publishing.
- Flowerdew, Lynne. (2005). An integration of corpus - based and genre - based approaches to text analysis in EAP / ESP: Countering criticisms against corpus-based methodologies. *English for specific purposes*, 24(3), 321-332.

- Flowerdew, Lynne. (2008). Corpora and context in professional writing. In Vijay. K. Bhatia, John. Flowerdew, Lynne and Rodney, Jones (Eds.), *Advances in Discourse Studies* (pp. 115-127). London: Routledge.
- Fulcher, Glenn and Davidson, Fred. (2007). *Language Testing and Assessment*. London and New York: Routledge.
- Garret, Peter and Shortall, Terry. (2002). Learners' evaluation of teacher-fronted and student-centred classroom activities. *Language Teaching Research*, 6(1), 25-57.
- Graves, Kathleen. (2001). A framework of course development processes. In David. R., Hall and Ann Hewings (Eds.), *Innovation in English language teaching: a reader*, (pp. 178-196). London: Routledge.
- Guilloteaux, Marie J., and Dörnyei, Zoltán. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly*, 42(1), 55-77.
- Härmälä, Marita. (2010). Linguistic, sociolinguistic, and pragmatic competence as criteria in assessing vocational language skills: the case of Finland. *Melbourne Papers in Language Testing*, 15(2), 1-43.
- Harmer, Jeremy. (2007). *The practice of English language teaching*. London: Pearson Education Limited.
- Hedge, Tricia. (2000). *Teaching and Learning in the Language Classroom*. Oxford: Oxford University Press.
- Hughes, Arthur. (2003). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Jiménez, Carlos. (2004). El Marco Europeo Común de Referencia para las Lenguas y la comprensión teórica del conocimiento del lenguaje: exploración de una normatividad flexible para emprender acciones educativas. *Revista FESC. Divulgación Científica Multidisciplinaria*, 4(14), 1-22.
- Kalaja, Paula and Ferreira, Ana María. (2003). *Beliefs about SLA: New Research Approaches*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Kaufman, Dorit. (2004). Constructivist issues in language learning and teaching. *Annual Review of Applied Linguistics*, 24(1), 303-319.
- Lantolf, James, P., and Poehner, Matthew, E. (2004). Dynamic Assessment: Bringing the past into the future, *Journal of Applied Linguistics*, 1, 49-74.
- Lee, David. (2007). Corpora and discourse analysis: New ways of doing old things. In Vijay K. Bhatia, John Flowerdew and Rodney Jones (Eds.), *Advances in Discourse Studies* (pp. 86-99). London: Routledge.

- Leech, Nancy L. and Onwuegbuzie, Anthony J. (2011). Beyond Constant Comparison Qualitative Data Analysis: Using NVivo. *School Psychology Quarterly*, 26(1), 70-84.
- Mella, Orlando. (2003). *Metodología Cualitativa en Ciencias Sociales y Educación*. Santiago: Editorial Primus.
- McKay, Sandra. (2003). Teaching English as an international language: the Chilean context. *ELT Journal*, 57(2), 139-148.
- Ministerio de Educación (MINEDUC). (2002). *Objetivos Fundamentales y Contenidos Mínimos Obligatorios de la Educación Básica*. Santiago, Chile: Ministerio de Educación, Gobierno de Chile.
- Ministerio de Educación (MINEDUC). (2009). *Fundamentos del Ajuste Curricular en el sector de Idioma Extranjero, Inglés*. Santiago, Chile: Ministerio de Educación, Gobierno de Chile.
- Nunan, David and Bailey, Kathleen. (2009). *Exploring Second Language Classroom Research: A Comprehensive Guide*. Boston: Heinle, Cengage Learning.
- Nunan, David. (2004). *Task-Based Language Teaching*. Cambridge, UK: Cambridge University Press.
- O'Dowd, Elizabeth. (2010). The development of linguistic complexity: A functional continuum. *Language Teaching*, 45(3), 329-346.
- Orafi, Senussi and Borg, Simon. (2009). Intentions and realities in implementing a communicative curriculum reform. *System* 37, 243-253.
- Pérez, Sergio. (2009). *Etnocentrismo y Discriminación en la Educación Chilena: Consecuencias en la identidad sociocultural de estudiantes mapuches de la Región de la Araucanía* (Unpublished doctoral thesis). Universidad de Extremadura, Spain.
- Richards, Jack. (2001). *Curriculum Development in Language Teaching*. Melbourne: Cambridge University Press.
- Scrivener, Jim. (2005). *Learning teaching: A guidebook for English Language Teachers*. London: London: Macmillan ELT.
- Ur, Penny. (1999). *A Course in Language Teaching*. Cambridge: Cambridge University Press.
- Wittgenstein, Ludwig. (1972). *Philosophical investigations*. London: Pearson Education Limited.
- Yu, Chong, Ho, Jannasch-Pennell, Angel and Di Gangi, Samuel. (2011). Compatibility between Text Mining and Qualitative Research in the Perspectives of Grounded Theory, Content Analysis, and Reliability. *The Qualitative Report*, 16(3), 730-744.