The pragmatic fossilization of discourse markers in non-native speakers of English

Jesús Romero Trillo*
Departamento de Filología Inglesa, Universidad Autónoma de Madrid-(Cantoblanco)
28049-Madrid, Spain

Received 11 January 2001; received in revised form 13 January 2002

Abstract

The present article describes the phenomenon of “Pragmatic Fossilization” as one of the main problems that non-native speakers of English face in their learning process. The article discusses its possible origins, its features, and its implications in the evolution of a second language. Based on a corpus-driven analysis of discourse markers, the paper focuses on the evolution of these pragmatic elements in native and non-native speakers of English—children and adults. The analysis will show the difference in the use of these elements between both groups of speakers, with the possible pragmatic failures that may arise in communication. The detailed study of this phenomenon, supported by statistical tests, proves the role of “Pragmatic Fossilization” as an element to consider in second language learning and teaching.

Keywords: Pragmatic fossilization; Discourse markers; Non-native speakers; Corpora

1. Introduction

The present article explores the use of pragmatic markers in the speech of native and non-native speakers of English at two different ages: children and adults. The aim is to see to what extent their exposure to spoken pragmatic information in a foreign language is sufficient to acquire pragmatic markers coherently. The study is based on the analysis of two different corpora—native speakers and learners—,
where the former will be used as baseline data for the interpretation of the non-native language production in both age groups.

The hypothesis is that learners of a foreign language follow what I will call a “binary track” in their linguistic development: the formal vs. the pragmatic track. The formal track relates to the grammatical and semantic rules that conform the competent use of a given language; the pragmatic track, on the other hand, relates to the social use of language in different contexts and registers.

Native speakers of a language would develop both tracks simultaneously by means of natural language contact, and thus would establish a mutual relationship between both communication tracks. Non-native learners of a language in a non-target language environment, however, would develop the formal and the pragmatic tracks through formal instruction. The difficulty, therefore, is that the pragmatic track, linked to the cognitive, affective, and socio-cultural meanings expressed by language forms, is difficult to implement in educational syllabuses. In fact, the development of pragmatic competence demands a (pseudo)-natural foreign language context that is often almost impossible to produce in formal education.¹

As a consequence, native speakers follow what I shall call a “function-to-form developmental process”, where the need to communicate precedes the use of a form, as Painter (1999) demonstrates; whereas non-native speakers follow a “form-to-function process”, based on the learning of certain items which are usually contextualized at different subsequent stages.

2. Pragmatic fossilization

In general terms, it can be argued that the form-to-function language development process that is characteristic of non-native speakers, is the key factor in the appearance of certain aspects of language fossilization. Traditionally, fossilization was identified with the persistence of formal (grammatical, semantic, phonological...) errors in non-native speakers (Selinker, 1972), although I believe that this phenomenon also applies to the pragmatic aspects of language development. In this article, I will address the phenomenon of what I shall call “Pragmatic Fossilization”, that can be defined as “the phenomenon by which a non-native speaker systematically uses certain forms inappropriately at the pragmatic level of communication”.

This approach leads us to one of the quandaries in foreign language learning (and teaching): what is pragmatically appropriate in a given situation? In other words, how can a language learner decide on the most appropriate element in a certain situation from a pragmatic perspective? Pragmatic competence does not consist of a series of rules that have to be observed in the communicative process, as it is the case with phonology and other aspects of language, but rather of a cline of appropriateness or acceptability.

Some linguistic approaches merge the formal and the pragmatic tracks and state that forms portray meaning and vice-versa. For instance, the systemic-functional

¹ See Kecskes and Papp (2000), ch.6 for an excellent account of the role of pragmatics in the development of an L2 in formal settings.
tradition, following Halliday (1975), considers both aspects of communication jointly and suggests that semiotic meaning is a way of saying and behaving in a language. Similarly, the cognitive tradition, as for example in Langacker (1990), also merges both tracks in order to explain linguistic communication in a comprehensive way.

Although the combination of the formal and pragmatic tracks might apply in the case of native speakers, I believe that non-native speakers find the formal-functional dichotomy problematic in their learning process because of the de-contextualized nature of teaching environments. Often, learners acquire a simplified and context-free register of the target language with no explicit relationship between form and function in most cases. Therefore, there is a need to investigate and teach the pragmatic function in relation to the cultural specificity of a language, ranging from the most transparent cases, such as lexicalized politeness phenomena, to the least obvious, such as the functions of discourse markers in conversation.

In my opinion, the study of such pragmatic fossilization can only be done using corpus-driven data, and would cover three different stages. The first stage involves the description of the pragmatic mechanisms of a language in order to study the differences and preferences in the realization of certain functions (Romero Trillo, 2002). The second stage relates to the specific pragmatic options that two or more languages offer in the realization of pragmatic functions (Romero Trillo, 1997, 2000), and the third stage would look at the pragmatic differences relative to different degrees of proficiency in a language. This will show the pragmatic competence of non-native learners of a language and its development throughout time (Romero and Llinares, forthcoming). This third stage can thus establish the cognitive and sociolinguistic factors that operate in language development.

During the learning process, non-native speakers are usually exposed to what has been called “semantic distance” between two languages (Hasan, 1984:193)—the principles that organize semiotic meanings in a language—and “cultural distance” (Kecskes and Papp, 2000:96)—the cultural and conceptual structure that models meaning in a language. However, non-native speakers are not exposed to what I shall call “pragmatic distance”: “the variants in the social, cognitive and contextual dimensions of linguistic communication that govern and systematize social relations in speech”. Accordingly, pragmatic distance would be an essential factor in communication as it models the relationship that exists between two speakers of a language (Kasper and Blum-Kulka, 1993).

In sum, pragmatic fossilization appears in the second language learner not because of a lack of competence in other linguistic areas such as lexis, grammar, etc., but because there is a delay in the presentation of the pragmatic variation that exists with respect to the way communication competence is acquired in the mother tongue.

3. Pragmatics and language development

Interactional approaches to language acquisition have followed two research paths that can be summarized as follows: the role of interaction in the development of grammar and lexis (Mackay, 1999); and the role of interaction in the development
of pragmatic competence (Cathcart, 1986). The latter approach has shown that pragmatic development (though not its verbalization) appears prior to the development of grammar, as it is shown in the fact that pre-linguistic infants can take turns, get the attention of others, direct that attention to objects of interest, make demands, etc., before they show any evidence of having control of grammatical structures. By way of illustration, Stern et al. (1975) showed that four-month-olds acquire turn-taking in their pre-linguistic interaction with their mothers. According to these authors, this principled interaction occurs in an orderly fashion, except when the mother intervenes trying to stop her baby from crying. However, when the child laughs and makes “nice” noises, the mother engages in a dialogic chorus with the child. As is obvious, this initial dyadic pattern is not repeated in formal L2 learning. Furthermore, the fact that the L2 learner has mastered turn-taking in his/her own native language is no guarantee for this in the other language.2

Other aspects of the acquisition of pragmatic meanings have been studied by Weeks (1971), who observed that different voice qualities signal different functions of the speech in young infants. For instance, soft speech appears when talking to oneself or when expressing futility; loud speech is used for correcting other people, scolding, or objecting; extra clear speech for correcting or avoiding misunderstanding; and fussy speech appears when one is undecided, or when hurrying to other more important parts of an utterance. High pitch is used for uncertainty or complaining; baby-talk pronunciations when talking about a younger child; and exaggerated intonation in story telling. This evidence suggests the way pragmatic competence is learnt, and its importance in language development from very early ages. Evidently, we need to look for a way of including this component in non-native teaching contexts, of course, taking into consideration the age of the learners.

4. A corpus study: pragmatic fossilisation and discourse markers in speech

The present study will compare the use of pragmatic markers in speech in native and non-native speakers of English at two different stages of development: children and adults. The aim is to see to what extent exposure to spoken pragmatic information in Spanish educational settings is sufficient to enable non-native speakers to use pragmatic markers appropriately. As mentioned above, the research procedure will consist of the analysis of different corpora in order to describe the patterns and the distribution of the use of discourse markers according to two variables: age and mother tongue.

The use of corpora for the analysis of spoken language is well established since the early 1980’s, and has given very fruitful results, as shown for example in Aijmer and Altenberg (1991). This subdiscipline of linguistics has enabled researchers to acquire sound methodological principles in the description of natural language (McEnery and Wilson, 1996; Oakes, 1998). In recent times, the compilation of corpora of natural language has begun to include the storage of non-native speech produced by learners.

2 As in, for example, the different rhythm of interruption in English and Spanish conversations.
of a language—especially of English. This new development enables teachers to assess the production and learning problems of students of a second or foreign language (Granger, 1998). In the present study, I will use both kinds of corpora in order to compare the development of certain discourse markers in the speech of native and non-native speakers of English. For this purpose, I selected data of two different ages: children (first graders) and adults. The data was collected as follows:

1. Native speakers:
   - Children: CHILDES corpus: Carterette & Jones Corpus, and Evans corpus
   - Adults: London-Lund corpus and Carterette & Jones corpus (psychology students)
2. Non-native speakers:
   - Children: Universidad Autónoma de Madrid (UAM)-Corpus: a section of pupils in a bilingual school in Madrid
   - Adults: UAM-Corpus: English philology students

The CHILDES corpus is a collection of data compiled by McWhinney and Snow (1998) with naturally occurring interaction of children from different social, linguistic, and educational backgrounds. Within CHILDES, I selected the Carterette and Jones Corpus, that consists of a collection of conversations of 54 first graders from two different schools of middle socio-economic level. The conversations reproduce exchanges between three children and a young adult who initiates the conversation. I also used the Evans corpus, a collection of 22 dyadic conversations between first graders at indoor play. Both corpora were chosen because they display the situational contexts, adult-child group interaction and child–child interaction that are also present in the non-native corpus under study, thus guaranteeing the comparability of both sets.

The adults’ section was taken from ten conversations of the London-Lund corpus, collected by Svartvik and Quirk, which reproduces interactions of 2 to 4 adults in a university environment (professors, students, secretaries, laboratory workers, etc.). The Carterette and Jones adults’ section consists of 24 conversations of triads of psychology students.

The non-native data comes from a section of the Universidad Autónoma de Madrid Corpus (UAM-Corpus). This corpus, still under development, is a longitudinal study of spoken English at six Madrid schools that teach English from preschool onwards. The corpus also includes the spoken production of students of English Philology at University. This part forms the Spanish section of the LINDSEI Project. For the present study, I selected the classroom interactions of first graders in a bilingual school in Madrid in different situations, and the speech pro-

---

3 I have selected this school because it probably represents one of the best educational environments to learn English as a foreign language in Spain. Likewise, the students of English Philology (in the 3rd and 4th year) are supposed to be the best sample of Spanish speakers of English.

4 For the present study, the CHILDES 1998 version has been used.

5 The LINDSEI Project directed by Prof. Granger consists of the compilation of a spoken corpus of non-native advanced learners of English in several countries.
duction of 3rd and 4th year English Philology students at the Universidad Autónoma de Madrid. The adult section of native and non-native speakers also share essential comparability features, since all speakers belong to the academic environment, and much of their talk concerns university life in its various facets.

Fig. 1 shows the sampling procedure for the study. Each box represents a section of the corpus, and the numbers inside indicate the total number of tokens in the interaction.

As mentioned above, the hypothesis is that pragmatic functions do not follow the same developmental pace as other linguistic layers, such as grammar or lexis. The analysis draws on previous studies that have examined the “pragmatic distance” of some functions in English and Spanish conversations (Romero Trillo, 1997, 2000). The study will concentrate on the realization of the pragmatic functions that I have found to be difficult in some way for Spanish non-native learners of English, because of the asymmetry between Spanish and English.

Specifically, I have concentrated on the use of discourse markers because they are elements that have no apparent meaning or grammatical ascription, are elusive to classification, but play a fundamental role in the pragmatic structure of interaction. The study of discourse markers has shown fruitful results since the mid-eighties. Schiffrin, (1987); Fraser, (1999) on the speech of adults; Kyrratzis and Ervin-Tripp (1999) on the development of discourse markers in 4 and 7 year olds in different activities; and Andersen et al. (1999) on the choice of markers in different language communities. However, in the literature there is a lack of studies on the acquisition of discourse markers by non-native learners of a language, and specifically, on the development and realization of the functions that these markers have in interaction.

In this line, in my previous work on discourse markers I have attempted to make functional descriptions of their appearance and of their variation in discourse, relying on statistical analyses of different corpora both in English and between this language and Spanish. From a theoretical perspective, Romero Trillo (2002) showed that discourse markers are elements that have undergone a process of discourse grammaticalization and have included in their semantic/grammatical meaning a pragmatic dimension that serves for interactive purposes. Here, I analyze the development and distribution of the canonical and the pragmatic use of the elements under study, looking for example at “well” as an adverb vs. “well” as an indicator of the beginning of a turn; or “you know” as a “sympathetic circular element” (i.e. the listener has the opportunity to express feedback but the floor returns to the speaker), or as an indicator of knowledge. Here are some examples:

- Canonical use:
  * then you’d probably do very well. (1.032 \carter\adult.cha 4)
  * ADU: Do you know Mister Adams (9.799carter\adult.cha 42)

- Pragmatic use:
  * JUL: I do-‘nt either.
  * JUL: so why are you asking me?
  * IVA: well why are you asking (414\evans\dyad09.cha24)
  * uh I go up there all the time and its really great you know (5.126 carter\adult.cha 22)
The specific discourse markers under study are “look”, “listen”, “you know”, “you see”, “I mean”, and “well”. “Look” and “listen” were selected because I had previously noted their different uses in adults’ speech in English and Spanish (Romero Trillo, 1997), and so I wanted to analyze the evolution in non-native speech compared to that of native speech. The other elements were selected because, of the totals of use of discourse markers in the London-Lund Corpus, they showed the highest percentage in native speakers, except for affirmative or negative elements (“yes, yeah, no”) and the hesitation marker “m”. Specifically, “you know” occurs 9.27%, “you see” 2.73%, “well” 10.33% and “I mean” 3.18% (Romero Trillo, 2002).

5. Analysis of the data

5.1. The native corpora

In order to establish comparative parameters for the use of discourse markers, I will first describe their distribution in native speakers with specific attention to the evolution from child to adult speech, which I use as the baseline for comparison.

In the analysis, I indicate the overall presence of an element and the number of times this element has a pragmatic function. With this information, I calculate what I call the “Index of Pragmatic Use” (hence IPU), that shows the degree of pragmatic use of an element in the corpus. This index is calculated as follows: number of pragmatic realizations, divided by the overall presence of the element. The IPU shows the degree of pragmatic specialization of a given item and is very helpful for stylistic and pedagogic purposes, because it shows the probability of an element being used as a pragmatic marker. The IPU can also be used for register analysis, since different styles of speech show different patterns of use of discourse markers, both in quantity and distribution.

Tables 1 and 2 show the results of children speech in the Carterette & Jones Corpus and Evans Corpus.

Looking at both corpora in combination, the analysis shows that the overall use of discourse markers in native children is the following: 300 pragmatic elements out of 66,463 words, i.e., 4.5%. In general terms it can be said that this figure is quite

<table>
<thead>
<tr>
<th>Element</th>
<th>Pragmatic function</th>
<th>Overall presence</th>
<th>IPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Look</td>
<td>2</td>
<td>6</td>
<td>33.3%</td>
</tr>
<tr>
<td>Listen</td>
<td>0</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>You know</td>
<td>43</td>
<td>86</td>
<td>53.4%</td>
</tr>
<tr>
<td>I mean</td>
<td>15</td>
<td>15</td>
<td>100%</td>
</tr>
<tr>
<td>Well</td>
<td>89</td>
<td>91</td>
<td>97.8%</td>
</tr>
<tr>
<td>You see</td>
<td>93</td>
<td>108</td>
<td>86.1%</td>
</tr>
</tbody>
</table>
low, which may indicate that most of the talk in children’s conversations is action-based (speech to obtain goods or services), not interaction-based. Therefore, children do not use discourse markers very much, since the main function of these elements is to build the interactive scaffolding of a conversation from a social perspective. These results also indicate the need for a categorization of discourse markers in two main groups: one, “involvement markers”: the elements that enhance the positive face of the interlocutor because they try to involve the listener in the thinking process of the speech (in the present study: “you know, you see, well, I mean”), and two, those elements that are more concerned with the operative process of the interaction, “operative markers”, which, as in the case of the attention-getting elements “look” and “listen”, try to make the conversation flow without any disruption (Romero Trillo, 1997).

The most interesting element to comment on is “well”. For example, in the Carterette and Jones corpus, 61 out of the 91 times it appears (67.8%), it is used to start a turn. The data also shows an interesting correlation, because “well” is significantly followed by the element “I”. This implies that children use this element to start a turn that is going to deal with their cognitive and social reality.

Regarding the adults, Tables 3 and 4 show the distribution in the use of pragmatic markers in the two corpora.

The results of the analysis show that there is an increase in the use of pragmatic markers in comparison with the children’s data: 1,393 elements out of 74,887 words, i.e., 18.6%. This figure indicates that native adults use the discourse markers under study around four times more than children do.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Results of the Evans Corpus—children: (28,512 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element</td>
<td>Pragmatic function</td>
</tr>
<tr>
<td>Look</td>
<td>27</td>
</tr>
<tr>
<td>Listen</td>
<td>0</td>
</tr>
<tr>
<td>You know</td>
<td>6</td>
</tr>
<tr>
<td>I mean</td>
<td>1</td>
</tr>
<tr>
<td>Well</td>
<td>24</td>
</tr>
<tr>
<td>You see</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Results of the London-Lund Corpus—adults: (50,000 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element</td>
<td>Pragmatic function</td>
</tr>
<tr>
<td>Look</td>
<td>11</td>
</tr>
<tr>
<td>Listen</td>
<td>1</td>
</tr>
<tr>
<td>You know</td>
<td>394</td>
</tr>
<tr>
<td>I mean</td>
<td>115</td>
</tr>
<tr>
<td>Well</td>
<td>439</td>
</tr>
<tr>
<td>You see</td>
<td>93</td>
</tr>
</tbody>
</table>
These differences are coherent with the more social nature of adult interaction where speech is not so much “action-based” and personal opinions are pragmatically embedded in the conversation. Examples: the increase of “I mean” to reformulate one’s statements, of “you know” to enhance the face of the addressee in what I call “sympathetic circularity function”, or the more frequent use of “well” to downgrade the personal opinion in the initiation of a turn.

In general terms, the data confirms that “involvement markers”, i.e.: “you know, I mean, well, you see”, are more frequent and appear mainly as pragmatic elements in natural speech (“you know” 90.9%; “I mean” 95%; “well” 87.4%, and “you see” 86.1%). On the contrary, “operative markers” i.e.: “look, listen”, have a lower percentage of use as pragmatic elements (“look” 35.4% and “listen” 33.3% -the latter with only one function).

5.2. The non-native corpora

Table 5 shows the results of the use of discourse markers in non-native children:

The first striking result is the increase in the presence of the “attention-getting” elements “look” and “listen” in comparison with the previous data. In the native corpora the children did not use the element “listen” to obtain the attention of the others, whereas in the non-native context “listen” appears 28 times, with an IPU of 73.7%. In general, non-native children seem to prefer this marker to the more polite “look” which, although extensively used as well, 24 times, has a lower IPU, 14.1%. This preference might be due to the fact that the Spanish direct translation of “listen”

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Results of the Carterette and Jones Corpus—adults: (24,887 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element</td>
<td>Pragmatic function</td>
</tr>
<tr>
<td>Look</td>
<td>2</td>
</tr>
<tr>
<td>Listen</td>
<td>0</td>
</tr>
<tr>
<td>You know</td>
<td>113</td>
</tr>
<tr>
<td>I mean</td>
<td>40</td>
</tr>
<tr>
<td>Well</td>
<td>182</td>
</tr>
<tr>
<td>You see</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Results of the UAM-Corpus—children: (35,536 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element</td>
<td>Pragmatic function</td>
</tr>
<tr>
<td>Look</td>
<td>24</td>
</tr>
<tr>
<td>Listen</td>
<td>28</td>
</tr>
<tr>
<td>You know</td>
<td>3</td>
</tr>
<tr>
<td>I mean</td>
<td>1</td>
</tr>
<tr>
<td>Well</td>
<td>22</td>
</tr>
<tr>
<td>You see</td>
<td>6</td>
</tr>
</tbody>
</table>
(“oye”) is not considered impolite and is indeed very frequent in Spanish conversations (Romero Trillo, 1997). The over-use of this marker confirms the non-naturalistic pragmatic input that foreign learners are exposed to at school, the only context where they will acquire the Foreign Language.

This high number of attention-getting markers correlates with the low use of all the other markers. For example, the use of the “well” is mainly confined to the evaluation of a task in the collocation “well done” (47.1%), and in only 11 instances (12.2%) the element “well” starts a turn -only once followed by the pronoun ‘I’. In fact, its IPU is 24.4%, compared with the 97.8% in the C&J and 96% in the Evans corpus.

Table 6 shows the results of the non-native adults.

Compared to the children’s data, here the markers (for example “you know”) increase in use, but much less so than in the native adult corpora. The case of “well” is again especially significant, because in only 53% of the times does it show a pragmatic use, as against the over 85% of use of this marker as a pragmatic device by both native children and adults.

As to “look” and “listen”, these markers have been completely eliminated from the corpus, although native speakers of English use these elements, especially “look”, to draw the attention of the addressee in casual conversation. It is very likely that non-native speakers identify the use of “look” and “listen” with impolite situations and, therefore, omit them in their dialogues.

In general terms, it can be said that the appearance of these markers in non-native adults is even more limited than in the native children corpora. In my opinion, this result is very worrying because, although the Spanish informants have a good grammatical level in English, the lack of these devices may prevent them from carrying out interactions efficiently, for example, being incapable of drawing the attention of the addressee in a way which is seen as polite, or of expressing personal opinions with the expected attenuating “well”.

5.3. Statistical analyses of the markers

In order to confirm the significance of these differences in the use of the markers between native and non-native speakers, and between children and adults, I carried out an ANOVA test for each marker under study. As is well known, this test will show if each of the independent variables, here: age and mother tongue, has a significant

<table>
<thead>
<tr>
<th>Element</th>
<th>Pragmatic function</th>
<th>Overall presence</th>
<th>IPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Look</td>
<td>0</td>
<td>16</td>
<td>0%</td>
</tr>
<tr>
<td>Listen</td>
<td>0</td>
<td>10</td>
<td>0%</td>
</tr>
<tr>
<td>You know</td>
<td>27</td>
<td>33</td>
<td>81.8%</td>
</tr>
<tr>
<td>I mean</td>
<td>26</td>
<td>26</td>
<td>100%</td>
</tr>
<tr>
<td>Well</td>
<td>86</td>
<td>162</td>
<td>53%</td>
</tr>
<tr>
<td>You see</td>
<td>2</td>
<td>20</td>
<td>10%</td>
</tr>
</tbody>
</table>
effect on the dependent variables, i.e., each of the markers. In order to check whether the two independent variables were related, I also carried out two-way ANOVA tests with each of the markers under study.

In Fig. 2 we can see that “I mean” is practically inexistent in children’s speech, both native and non-native, but that native adults use this element in their argumentation, though it is absent from non-native talk. The statistical tests gave the following significant results: for the language variable ($F = 5.69; P = 0.02$), for the age variable ($F = 55.37; p < 0.00001$), and for the interaction of both ($F = 5.71; P = 0.02$). This means that this difference is statistically significant.

Fig. 3 presents an opposite situation, the complete absence in the use of the face-threatening element “listen” in native speakers, but the very high use of this element in the non-native children corpus, followed by a decrease in the use of the element by adults. The three ANOVA tests showed the same very significant result of $F = 9.7; P = 0.003$.

The case of “look” (Fig. 4) presents a much more balanced situation in the two groups of speakers: native and non-native. In fact, from a statistical point of view the only significant difference was not in the language variable or in the interaction, but in the age variable. The test shows that there is a significant decline in the use of the marker in adult conversations ($F = 5.48; P = 0.02$). This is the only element that non-native speakers seem to use appropriately, compared to its use by native speakers.
Fig. 5 shows the significant difference in the use and evolution of the element “well”. The graph shows the difference in the use of the marker between children and adults ($F=51.17; P<0.00001$), as well as between natives and non-natives ($F=9.98; P=0.002$), with also a significant interaction in the two-way ANOVA test ($F=9.56; P=0.03$) of both variables. These results indicate a failure in the pragmatic competence of non-native speakers, as indicated above.

The pattern of the graph in Fig. 6 shows a similar evolution in the marker “you know” in native adults. The statistical analyses give very significant results for the two variables: language has a value of $F=18.15 (P=0.0009)$; age has a value of $F=34.68 (P<0.00001)$, and their interaction $F=17.28 (P=0.0001)$. These values are the most significant ones of the whole series of markers under study here. Again, the pragmatic load of the element seems to be missing in non-native adult speech.

The last graph shows that the use of the element “you see” is almost non-existent in non-native speech, with the same amount of use in children as in adults. The statistical values of the mother tongue variable ($F=12.22; P=0.0009$) and of the age variable ($F=12.09; P=0.001$) are very significant. The significant result of the two-way ANOVA test ($F=12.42; P=0.0003$) reaffirms the urgent need to implement the teaching of this element to foreign speakers.
The data in the present analysis has shown that native speakers increase their use of “involvement markers” as they grow up, whereas non-native speakers do not master this group of pragmatic elements and fossilize with a reduced proportion of use. On the other hand, non-native speakers are able to use “operative markers” in a way comparable to their use by native speakers. This may be because “operative markers” are linked to the mechanics of interaction, whereas “involvement markers” are more concerned with the articulation of the argumentative process and the social and cognitive relationship between the speaker and the addressee.

6. Conclusions

In this section, I first highlight the importance of addressing pragmatic questions using a corpus driven approach. The availability of a large volume of data with different characteristics, as used in the present study, guarantees the objectivity of the results and opens paths for future research where the variables of the corpora may be different.

Regarding the specific study of pragmatic fossilization:
1. The first conclusion of this study is that there is a different rate of development for the grammatical and the pragmatic aspects of language in L2. This can be observed in the linguistic production of proficient non-native speakers of English who do not show a competent use of the pragmatic functions needed in casual conversation. This can be due to the non-natural teaching environment where non-native speakers learn the L2. In fact, the quantitative analyses indicate that native and non-native children show a similar pattern in the use of the markers (except in the case of “listen”) but that when they grow up, non-native speakers fail to acquire the appropriate markers that scaffold adults’ speech.

2. Children’s conversation is action-based rather than interaction-based, as the choice of markers demonstrates. Adults’ conversation, on the contrary, is interaction-based and demands a competent use of involvement markers rather than operative markers. The lack of this competence leads to pragmatic fossilization and, possibly, to communicative failure in many cases.

3. Still, if pragmatic functions were introduced in the teaching process, foreign children might pick up the pragmatic value of linguistic elements in the same way as native children, as the quantitative analyses indicate. This knowledge is neglected in the curriculum and, thus, the use of pragmatic markers becomes fossilized both in the quantity and the diversity of elements used. To some extent, it can be said that non-native speakers are deprived of many pragmatic resources in their L2 learning process.

The present paper has offered a corpus-driven comparison in the use of discourse markers in native and non-native children and adults and has demonstrated the urgent need to bring the consistent teaching of pragmatic markers to language instruction. This teaching has to be based on sound research studies that categorize and describe the use of these pragmatic elements on the basis of the Index of Pragmatic Use and of other mathematical and statistical methods (Romero Trillo, 2002). To sum up, I am convinced of the need to investigate the development of pragmatic markers in speech, in order to monitor pragmatic competence and pragmatic fossilization in non-native speakers.

References


Romero Trillo, J., Llinares García, A. & Fernández Agüero, M. (in progress) UAM-Corpus. i The LINDSEI Project, directed by Prof. Sylvianne Granger (Université Catholique de Louvain), consists in the compilation of a spoken corpus of non-native advanced learners of English in many countries of the world.


