The Database of Cartoons and Commercials as Visually-Oriented Language Input: Motivations, The Procedure and Methodological Issues*

Pınar İbe Akcan¹, Umut Ufuk Demirhan²

¹, ² Mersin Üniversitesi, Fen-Edebiyat Fakültesi, İngiliz Dili ve Edebiyatı Bölümü, 33343 Çiftlikköy/Mersin
¹ pinaribe@gmail.com, ² umutufuk@gmail.com

(Received 18 February 2016; accepted 30 March 2016)

ABSTRACT: Cartoons and commercials on television may be important resources for the acquisition and learning processes of children in both linguistic and non-linguistic domains. However, there are not adequately many studies in the linguistics literature about them, particularly emphasizing their importance as a form of visually oriented language input in the language acquisition process. The aim of the current study is to introduce a database of cartoons and commercials in Turkish, presenting its major motivations, its construction procedure and some of the basic issues that were experienced during the construction process. The paper aims to make methodological contributions for further studies that have similar objectives.

Keywords: language use, language acquisition, cartoons, commercials

* The project (Cartoons and Commercials as Visually Oriented Language Input) is supported by the Mersin University Scientific Research Projects Fund with the Project Number 2015-AP2-1238.

http://dx.doi.org/10.18492/dad.06656

Dilbilim Araştırmaları Dergisi, 2016/1, 1-15.
© 2016 Boğaziçi Üniversitesi Yayınevi, İstanbul.
1 Introduction: Why Cartoons?

While the child acquisition literature is quite rich in terms of the language input sources in the language acquisition process (Kempe and Brooks, 2001; Dominey and Dodane, 2004; Buttery and Korhonen, 2006; Murphy and Jones, 2008; Rowe, 2008; Weisleider and Waxman, 2010; Veneziano and Parisse, 2010 and many others) and, the pedagogical literature frequently mentions the use of cartoons as language teaching tools (Fleming, 1964; Brinton, 1991; Clark, 2000; Doring, 2002; Gillmore, 2007; Yaman, 2010; Bahrani, 2012), studies specifically focusing on cartoons and commercials on television as a form of language input are rare to find. In fact, cartoons and commercials on television can be highly effective in the acquisition and learning processes of children in the linguistic and non-linguistic domains. Studies on cartoons and commercials in the literature often focus on the media effects on children from either physical and cognitive perspectives (Bee and Boyd, 2000; Van Evra, 2004; Oruç and Tecim 2011; Özatar and Koç 2012). Regarding the cognitive domain, the literature is rich in dealing with issues such as the media effects on children’s development and the hidden messages that cartoons impose (Aitken, 1986; Dodd 1992; Barbara, 1993; Middleton and Vanterpool, 1999; Bee and Boyd, 2000; Perse, 2001; Van Evra, 2004; Stagg, 2005; Mitchell, 2014). There are few studies in Turkish from the same perspective (Oruç and Tecim, 2011; Özatar and Koç, 2012) focusing on the effects of cartoons on children’s behaviors and personality development. To make the picture complete, we should mention that physical effects are analyzed in relation to children’s playing behavior, tendency to violence, aggressive behavior, hyperactivity, eating behavior, addictive behaviors and so on (Puzgles and Anderson, 1979; Watkins, 1980; Roedder, 1981; Singer, 1998).

From the perspectives of language acquisition and language learning, cartoons and commercials on television are an inseparable part of the experience of every child since a considerable amount of language input enters into the children’s storage of grammar and lexicon through such media. Cartoons can be expected to make a positive cognitive and psychological contribution to each process. On the other hand, both cartoons and commercials on television may have negative effects on children, particularly in terms of their socialization and the development of inferencing and critical thinking processes.

From the language teaching perspective, films, cartoons and other audio-visual instruments are widely used as authentic language teaching materials in first and second language teaching settings. The pedagogical value of audio-

---

1 Throughout the article, the terms cartoon and commercials refer to those on exclusively cartoon broadcasting TV channels.
visual materials becomes clear especially when one wants to improve the listening and vocabulary skills. It is clear that audio-visual materials will be more widely used in language teaching in the future to boost motivation, to increase the attention span, and to improve memory. Cartoons and commercials on television are very important in this respect because children are exposed to verbal and nonverbal, direct and indirect messages in and outside home or school from cartoons and commercials. But cartoons and commercials on television have been taken into consideration (especially in child psychology) mostly in terms of the negative messages that they implicitly impose (Aitken, 1986; Mitchell, 1995; Beron, 1993). A review of the Turkish literature also shows that there exist a number of studies critically analyzing the textbooks and other kind of written materials targeted to children (i.e. children’s literature and a multitude of recommendatory textbook criticisms like Çetinkaya and Uzun, 2011; Aksan and Uçar, 2012; İbe Akcan, 2012; 2014; Sever, 2013 to mention a few). However, to the best of our knowledge, there are not any studies particularly in Turkish related to the cartoons and commercials on television analyzed with respect to their linguistic or audio-visual role as language inputs. In a similar way, there are no studies in teaching Turkish as a first language dealing with the effects of cartoons and commercials on television. The Database of Cartoons and Commercials aims to provide a basis for such studies.

The outline of the rest of the study is as follows. Section 2 briefly presents the aims of the study. Section 3 details the transcription issues and the database construction procedure – the identification of cartoons and commercials, sampling and labeling, classification. It also provides the transcription preferences of the study and proposes a simple coding manual for further studies. Finally, Section 4 draws some conclusions and summarizes the study.

2 Aims of the Study

The Cartoons and Commercials Database Project aims to answer the following broad question: what is the nature of the visually oriented language input directed to Turkish children between ages over 6 and under 6? The linguistic analyses on the database are aimed to cover three major aspects of the language: the syntactic properties and regularities (dominant syntactic structures), the semantic properties and regularities (dominant semantic domains and sub-domains), and the pragmatic properties and regularities (hidden messages). In addition to these major linguistic aspects, the database may contribute to other linguistic issues (e.g. discursive and psycholinguistic aspects) leading to psycholinguistic analyses. All in all, the purpose of the current study is to introduce the Turkish Database of Cartoons and
Commercials, presenting its major motivations, the construction procedure and some of the methodological issues that have shown up. The ultimate aim is to contribute to studies that have similar objectives.

3 Related Issues and Database Construction Procedure

This section details some of the issues in transcription and the construction procedure of the database in terms of its substeps.

3.1 Issues in Transcription

As it is well-known, transcription is an important tool for language documentation; it is also a difficult process which requires smart decision making before, during and after the process itself. This section reviews several issues in the transcription of (spoken) language we found useful while constructing our database.

In manual human transcription processes, first and foremost, the transcription tools and the gathered data need to be standardized. Given the standardization, operability and sustainability issues, the construction of language databases is a time consuming and labor intensive activity, as Ruhi (2013) points out. The annotation of spoken corpus is also a difficult but a necessary process in understanding the discourse structure as Zeyrek, Turan and Bozşahin (2009) indicate. Valuable resources have been produced for spoken Turkish with the stated purposes (Ruhi, Hatipoğlu, Eröz Tuga and Karadaş, 2010; Ruhi, 2013) and there exist model proposals for the transcription of the spoken language data (Uçar, 2014). These resources have been informative for our purposes.

According to O’Connel and Kowal’s (1999), an analysis of the existing transcription systems that encode verbal, prosodic, paralinguistic and extra-linguistic features of spoken discourse reveal five issues within the current practice which are; options (related to the purposes of the researcher in his mind), lexical integrity (related to the legibility of the transcribed lexical items), notation (related to the use of the individual characters and their purpose of use), conventional typography (related to the adaptation of typographical symbols) and measurement (related to the bias of transcribers in the notation of measured features) (pp. 107-109;113-115).

They pose five questions which most researchers dealing with transcription issues have to reflect on (pp. 115-117). We provide them below and in parentheses, indicate how we used their guidelines.

1. How much should be encoded in a transcription system? (The answer is to be determined by the researchers’ selection.)
2. How should these features be encoded? (The corpus developer should adopt one system of transcription or develop one that fits the purpose of the study.)
3. How should the phenomena encoded be operationalized? (The result should be accessible to researchers.)
4. For whom is a given transcript intended? (The transcripts are tools for further analysis so transcribers should avoid interfering with the plain text.)
5. What then is to be said of standardization? (To the best of our knowledge, there is no agreement among researchers regarding standardization of notation.)

O’Connel and Kowal’s (1999) focus on two issues is noteworthy; the full adaptation of an available annotation system for cartoons and commercials is almost impossible since one would be restricted by a) the purpose of the study and b) the researchers’ selection.

Dressler and Kreuz (2000) propose six design principles for transcription systems which are; specificity, universality, consensus, transparency, parsimony, conventionality and expansibility (pp. 27-29). They group transcription conventions into the following categories (pp. 29-34):

1. Intonation (rising, falling, continuing intonation, intonational contours etc.)
2. Temporal features (pause, short untimed pause, slow, rapid and overlapping speech, lengthened syllable, word cutoff, latched talk, backchannel etc.)
3. Intensity (soft speech, loud speech etc.)
4. Breathing (audible breathing, in-breath, out-breath etc.)
5. Transcribers’ comments (paralinguistic behavior, unclear or unintelligible speech etc.)

Most of Dressler and Kreuz’s (2000) transcription categories do not seem to work for cartoons and commercials for the obvious reason that speakers are not human beings. Hence, the existing transcription systems are not likely to fit cartoons and commercials unless necessary modifications are implemented. We would like to stress that transcription issues may confuse a researcher building a database of cartoons and commercials. The reason is that the data to be transcribed are different from the ordinary spoken language, not only because the speakers are not human but also the characters’ speech does not reflect the normal speech of a human being; therefore, a set of different criteria will be needed for transcription.
3.2 The Database Construction Procedure

The database construction procedure mainly covered four steps. The pre-study step consisted of the identification of cartoons and commercials which would be included in the database. The second step was the constitution of the sample out of the universe (i.e. the gathering of cartoons from which a representative sample of such media is taken for the analysis) and the labeling. The third step involved the classification of the cartoons in terms of their original language (i.e. whether the cartoon or the commercial is originally Turkish or dubbed) and classification in terms of the target age groups (i.e targeted under 6 or over 6). The ultimate step was to analyze the data linguistically (syntactically, semantically and pragmatically). In sum, the major steps followed during the data gathering stage were as follows:

1. Identification of the exclusively cartoon broadcasting TV channels (by means of a web search) and identification of the most popular cartoons (by analyzing TV audience measurements on the web) and frequently broadcasted commercials,
2. The sampling of the cartoons and commercials (using the Maximum Variation Sampling Method),
3. The classification of the sample cartoons with respect to their originality (i.e. whether they are originally Turkish or dubbed) and with respect to the age group of the target audience (under 6/over 6),
4. The transcription and digitization of the classified cartoons or the transcription and the digitization procedures.

In what follows, steps 1-4 above are detailed. Then, a coding manual is presented.

3.2.1 Identifying cartoons and commercials

In the first phase, the open-access web sites on the internet, such as YouTube, Dailymotion and Vimeo were searched to identify the cartoons and commercials in the database. In addition to the above mentioned video channels, the official web sites of the TV channels were also used as a resource. By using Google Chrome’s “Video Downloader” plugin, the selected videos were downloaded in order to be able to work offline.

The broadcast stream of a TV channel is an ever-changing list, so the identification of a time period was necessary for a steady list. At this point, the broadcast streams of the TV channels were analyzed at a specified time point. The chosen interval was the summertime (June, July, August 2015) streams of cartoon broadcasting channels and the most popular ones were identified to be included in the sample. In the second phase, in order to
To compile a representative sample, TV audience measurements available on the web were overviewed without a detailed examination of their validity, assuming that a choice of a ten randomly selected scales retrieved by a google search was enough for evaluation.

Only two groups of commercials – food and toys were included in the database, based on the observation that their frequency of overrides the others.

### 3.2.2 Sampling and labeling of cartoons and commercials

In the second phase, a sample was built up from the universe of pooled cartoons and commercials using the Maximum Variation Sampling Method. This is a sampling method that allows the researcher to investigate the particularities of each sample as well as the shared patterns within the pool. It involves purposefully choosing a wide range of cases to get variation on different dimensions of the items to be analyzed (Patton, 1990, p.172).

In order to organize the data in a structured way, the markup information was added to each file. This is the coding or labeling procedure before the transcription of the media. Technically, a markup file or a file ID allows one to view the channel, the name, the episode number, the time period when the cartoon or commercial was cast, whether is originally Turkish or dubbed, and the target age. In this phase, the steps were formed as follows:

- The sampling of cartoons and commercials (Maximum Variation Sampling Method) (Table 1).
- The downloading, labeling or coding the sample media for convenient storage (Tables 2 and 3).

#### Table 1. Study sample from the universe of cartoons

<table>
<thead>
<tr>
<th>Number of Channels</th>
<th>Number of Cartoons (Selected of Total)</th>
<th>Number of Episodes (Cartoon X Episode= Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cartoon Network</td>
<td>6 of 26</td>
<td>6x3=18</td>
</tr>
<tr>
<td>Minika Çocuk</td>
<td>4 of 16</td>
<td>4x3=12</td>
</tr>
<tr>
<td>Minika Go</td>
<td>4 of 18</td>
<td>4x3=12</td>
</tr>
<tr>
<td>Planet Çocuk</td>
<td>2 of 6</td>
<td>2x3=6</td>
</tr>
<tr>
<td>TRT Çocuk</td>
<td>5 of 20</td>
<td>5x3=15</td>
</tr>
<tr>
<td>Yumurcak TV</td>
<td>5 of 19</td>
<td>5x3=15</td>
</tr>
<tr>
<td><strong>Total:</strong> 6</td>
<td><strong>Total:</strong> 26</td>
<td><strong>Total:</strong> 78</td>
</tr>
</tbody>
</table>

A markup template was designed for the sampled cartoons, as Table 2 and Table 3 display:
Table 2. Markup template for the cartoons

<table>
<thead>
<tr>
<th>Channel</th>
<th>Name</th>
<th>Episode</th>
<th>Period</th>
<th>Originality</th>
<th>Target</th>
<th>Different Speakers</th>
<th>After</th>
<th>During</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Original episode number)</td>
<td>(Total mns.)</td>
<td>(Turkish/Dubbed)</td>
<td>(T/D)</td>
<td>Age</td>
<td>(Under 6/Over 6)</td>
<td>(0/1)</td>
<td>Gender</td>
<td>(Male/Female/Other)</td>
</tr>
<tr>
<td>Cartoon</td>
<td>Transf.</td>
<td>01</td>
<td>10:12</td>
<td>T</td>
<td>1</td>
<td>Within the text</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. A sample ID for the cartoons

- Cartoon ID
  - Cartoon (C1)
  - Channel (Max. 2 characters)
  - Capitals of the name of the cartoon (Max. 3 characters)
  - Episode number (Max. 2 characters)
  - Originality
  - Target age group
  - Period

C1CNTRA10:12D1 (Sample ID)

For later processing phases and linguistic analyses (which are performed via AntConc (Anthony, 2016))², the files were saved with their unique IDs in order to track and identify them. For example, the ID C1CNTRA10:12D1 contains the contextual information below:

C1: The rank number of the video in cartoons
CN: The name of the channel (here, it is Cartoon Network)
TRA: The name of the cartoon (here, it is Transformers)

² AntConc (version w3.4.4, 2016) supports working on multiple text files by means of its built-in tools such as the Concordance, the Clusters/N-Grams/Collocates and the Word List. The Concordance tool shows how a linguistic structure is used, the File View tool is used for investigating the results in more detail, especially for pragmatic analyses, the Clusters/N-Grams/Collocates tools are used to analyze common expressions in the texts and collocational behaviors of the linguistic units such as word+word, word+phrase, phrase+word combinations; and lastly, the Word List tool is used for generating and counting all word types and their observed frequencies in the texts.
The length of the video
D: Whether it is originally Turkish or dubbed (here, it is dubbed)
I: Target age group (here, it is over 6)

The same data coding procedure was also used for the commercials as shown in Table 4:

Table 4. Study sample from the universe of commercials

<table>
<thead>
<tr>
<th>Channels</th>
<th>Number of Commercials</th>
<th>Number of Toys Commercials</th>
<th>Number of Food Commercials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cartoon Network</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Minika Çocuk</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Minika Go</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Planet Çocuk</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>TRT Çocuk</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Yumurçak TV</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>6</strong></td>
<td><strong>Total: 30</strong></td>
<td><strong>Total: 30</strong></td>
</tr>
</tbody>
</table>

A markup template was also designed for the commercials as Table 5 displays:

Table 5. The markup template and a sample ID for the commercials

<table>
<thead>
<tr>
<th>Commercial ID</th>
<th>Channel</th>
<th>Food/Toys</th>
<th>Period</th>
<th>Target Age Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Commercial (C2)</td>
<td>Channel</td>
<td>(F / T)</td>
<td>(Total minutes)</td>
<td>(Under 6/ Over 6)</td>
</tr>
<tr>
<td>• Channel (Max.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Food/Toys</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Target age group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Period</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2CNT01:101</td>
<td>Cartoon Network</td>
<td>T</td>
<td>01:10</td>
<td>1</td>
</tr>
</tbody>
</table>

The labeling of the media was performed before the transcription, except for the during-and-after information (cf. Table 2) supplied by the transcribers. After controlling the before-, during- and after-information supplied by the transcribers, the operation was approved as finalized.

3.2.3 Classification of cartoons

In the third step, sample cartoons were classified with respect to their original language (Turkish/dubbed) and the age group of the target audience. The underlying assumption is that being originally Turkish or dubbed will be of
importance in future analyses since we expect that the language use in each case will differ in certain ways.

The cut-off point among the target age groups of cartoons was not easy to specify since the signs (i.e. smart signs, which means the symbols that indicate the target audience appearing before a telecast) are either not available (the available ones mark only the audience at ages 7+, 13+, 18+, and the general audience), or only the nature of the production, e.g. whether it involves violence or horror is marked. Thus, a rough division between the pre-school age (under 6) and school age (over 6) was found sufficient for the current purposes.

3.2.4 Transcription preferences

The issue of why the transcription in the current study was performed manually instead of making use of a computerized transcription system is that even the best speech recognition softwares make certain errors. When there is external noise, music, overlaps among the speakers, background speech or some other sounds in the video files, the amount of errors will increase. For example, one of the disadvantages is that even natural human voice has lots of varieties (such as the non-standard speech forms, dialects, accents etc.) which require some other decision making processes on the part of the researcher. (Human voice was not codified and animated movies were not included in our database). Gender identification is another important handicap for speech-synthesis or speech-recognition softwares, which is one of the codified features in our database. All these seem to result in a failure of the applications’ measure of performance. The choice of manual human transcription over the computerized ones is one of the important issues on its own sake which is beyond the scope anyway.

The assumption in the present study is that the transcription is not simply the written form of spoken language but something more. It is the reconstruction of a text by another writer and a representation of the conceptual world of language users. However, our aim is neither a conceptual nor a discursive analysis of the speakers so, a more neutral stance of the transcriber is needed in the current study (in spite of the fact that no analysis including a human researcher is expected to be fully objective). The study is also not a multi-modal or a multi-layered analysis for the above mentioned obvious reasons (such as the impossibility of the annotation of body language or other paralinguistic factors). Therefore, the following basic steps were regarded as sufficient:
• Constructing a simple coding system which includes only the features to be analyzed,
• The use of human transcribers with necessary skills and attention.

3.3 The Transcription Manual

In the transcription manual, an English based codification was used; the dates, numbers and other numerical data (c.f. Table 6) were coded with Arabic numerals. The details for the transcribers are indicated in Tables 6 and 7:

Table 6. Coding details for the transcribers: What to transcribe?

<table>
<thead>
<tr>
<th>When to transcribe?</th>
<th>What to transcribe?</th>
<th>The code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>Channel</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Capitals of the name of the cartoon</td>
<td>A (Max. 2)</td>
</tr>
<tr>
<td></td>
<td>Episode number</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Period</td>
<td>00</td>
</tr>
<tr>
<td></td>
<td>Originality</td>
<td>T/D</td>
</tr>
<tr>
<td>During</td>
<td>Different speakers</td>
<td>A, B, C</td>
</tr>
<tr>
<td></td>
<td>Gender of the speaker (Male, Female, Other)</td>
<td>M, F, O</td>
</tr>
<tr>
<td>After</td>
<td>Target age group</td>
<td>-6/+6 (0/1)</td>
</tr>
</tbody>
</table>

Table 7. Coding details for the transcribers: What not to transcribe?

<table>
<thead>
<tr>
<th>What not to transcribe?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No annotation of physical environmental features (the setting information)</td>
</tr>
<tr>
<td>No annotation of supra-segmental/prosodic/paralinguistic features (duration, pitch, loudness, intonation, slowness/quickness, voice quality, tone of voice, breathing, sighing, other produced sounds (linguistic/non-linguistic), laughing/crying)</td>
</tr>
<tr>
<td>No annotation of overlapped/interrupted speech</td>
</tr>
<tr>
<td>No annotation of outside voices</td>
</tr>
<tr>
<td>No annotation of body language (linguistic/non-linguistic: nods, eyebrow movements)</td>
</tr>
<tr>
<td>No annotation of exclamations and discourse markers</td>
</tr>
<tr>
<td>No standardization of spoken language</td>
</tr>
</tbody>
</table>

A form including the transcription criteria was given to the transcribers. A free field was included in the forms so that the transcribers could input notes regarding the problems they encountered or the observations they wanted to report.
Using the designed simple notation system in Table 6 and Table 7 (which only consist of the features to be analyzed in the project), cartoons and commercials were transcribed by the volunteered human transcribers who were employed for a fee for per-word for the task. The demographics of the transcribers are indicated in Table 8 and Table 9:

**Table 8. The demographics of the cartoon transcribers**

| Size: 21 transcribers |
| Distribution: Sophomore:12 transcribers, Senior: 9 transcribers |
| Composition: |
| Occupation: University students |
| Age: Between 19 and 30 |
| Gender: 18 females 3 males |
| Native language: Turkish |
| Foreign language: English |

**Table 9. The demographics of the commercial transcribers**

| Size: 12 transcribers |
| Distribution: Senior:12 transcribers |
| Composition: |
| Occupation: University students |
| Age: Between 23 and 30 |
| Gender: 10 females 2 males |
| Native language: Turkish |
| Foreign language: English |

A time constraint was not placed on the transcription task. The how-to-tips of the application to be used (which is the Notepad++)\(^3\) were given along with the transcription manuals. The transcribers were also given oral instructions plus an online consultation possibility. Once the transcribers watched a specific video, they concurrently transcribed what they heard via Notepad++ and saved the files with the “.txt” file extension.

When all the transcribed texts have been created using the text formatting standards presented to the transcribers, the texts were re-checked for possible spelling and punctuation errors. So far, the transcribed texts have been

---

\(^3\) The Notepad++ is a relatively uncomplicated and a user friendly application very similar to the Windows’ built-in application Notepad. Technically, unlike Microsoft’s built-in application Notepad, Notepad++ offers both spell checking and a large variety of text encoding formats such as Unicode and UTF-8. (Using the UTF-8 encoding standard is a must for languages like Turkish, since most of such applications are designed for the recognition of the Roman Alphabet and present a very limited language encoding options to the users).
analyzed via AntConc (version w3.4.4, 2016) which is a freeware text analysis toolkit perfectly working on Turkish texts (as corpus linguists are familiar with). The frequencies have been quantified and tabulated. We are currently analyzing the data syntactically, semantically and pragmatically.

4 Conclusion

In this paper it was argued that there is a need for the linguistic analyses of cartoons and commercials directed to children of different age groups. Constructing the databases of cartoons and commercials is a new but a necessary step to carry out linguistic analyses. The current study described the transcription procedures and the coding principles of the process. Needless to say, since each transcription manual depends on the researcher’s specific purpose(s), the current brief account of the procedure, the technical details and methodological issues encountered with during the process are just hoped to guide prospective researchers building or aiming to build such particular-purpose, small scale-databases.

References


