Education Quarterly Reviews

The online version of this article can be found at:
https://www.asianinstituteofresearch.org/

The Asian Institute of Research *Education Quarterly Reviews* is a peer-reviewed International Journal. The journal covers scholarly articles in the fields of education, linguistics, literature, educational theory, research, and methodologies, curriculum, elementary and secondary education, higher education, foreign language education, teaching and learning, teacher education, education of special groups, and other fields of study related to education. As the journal is Open Access, it ensures high visibility and the increase of citations for all research articles published. The *Education Quarterly Reviews* aims to facilitate scholarly work on recent theoretical and practical aspects of education.

ASIAN INSTITUTE OF RESEARCH
Connecting Scholars Worldwide

Common and Productive Morphemes in Language Acquisition: A Corpus-Based Study on Children's Books

Elif Sütçü[1]

[1] School of Foreign Languages, Başkent University, Ankara, Türkiye

Correspondence: Elif Sütçü, School of Foreign Languages, Başkent University, Ankara, Türkiye.
E-mail: elif2708@gmail.com

Abstract

This study is a corpus study on morphological features of children's books in Turkish. The purpose of the study is to determine derivational and inflectional suffixes in Turkish children books and to identify the common and productive morphemes used in child literature in the light of the information obtained; hence, reach clues about determining the priorities in the process of vocabulary acquisition in children. To achieve this, different literature examples written for children were examined, a corpus was formed by analyzing the derivational and inflectional suffixes, the frequency of use of categories and sub-categories was determined and a comparison between categories was made. The results yielded that inflectional suffixes were used more frequently than derivational suffixes, which could be due to less frequent use of derivational suffixes than inflectional suffixes. Also, derivational suffixes could be less productive in language acquisition process and therefore acquired late.

Keywords: Language Acquisition, Inflectional and Derivational Morphemes, Corpus

1. Introduction

Language, which we use to express our feelings and thoughts, is a human-specific phenomenon that distinguishes us from other living things. Although language has a more complex structure than many of the skills we acquire, every child acquires language at the end of a similar time and process (Macwhinney, 2003). In order to understand how this process takes place, the child's language (word) production is observed from infancy, and possible mental processes are tried to be determined.

One of the important factors in determining how language acquisition takes place is the frequency of use of words and units. In the acquisition process, if the child hears a word or a linguistic unit more and processes it into memory, he or she will have acquired that structure earlier. Therefore, the child pays attention to how often he hears morphemes and how many different structures these morphemes are used in (Clark, 2009). Accordingly, morphemic frequency analysis has an important place in observing word acquisition and finding clues about how it occurs. While conducting this analysis, written language can also be made use of besides the spoken language that children are naturally exposed to. Since early childhood books are written for children who have just acquired or are acquiring a language, they can provide us with information in areas such as the order of acquisition of

Turkish and the primary vocabulary. In this study, the books belonging to this period were examined, their morphological structures were analyzed, and the relationship between the affixes and the acquisition of Turkish was tried to be determined. Since no such studies in Turkish in the literature could be reached at the time, the aim of this study is (a) to create a small corpus by analyzing randomly selected books from early childhood literature; (b) to conduct a detailed morphological analysis and to determine derivational and inflectional suffixes in Turkish early childhood books and (c) to determine the common and productive vocabulary used in early childhood literature and in language acquisition process respectively.

In this paper, firstly basic notions and main studies related to morphological frequency and productivity will be mentioned. Then two main groups of suffixes in Turkish, namely derivational and inflectional suffixes will be presented in detail. After that some of the main corpus studies in Turkey will be discussed.

1.1 Frequency and Productivity

The concept of frequency and productivity is one of the key concepts of morphology. Thus, in this section some basic notions will be presented and the relation between frequency and productivity will be discussed.

The notion of frequency we mention in this study is "token frequency." Token frequency is concerned with the actual number of occurrences within the morphological category under consideration in a given text. This means that "repetitions of the same word count as separate items in the token frequency total" (Bauer 2001: 47). Accordingly, in this paper, objective frequency which is the actual number of times a word occurs in a specific written or transcribed spoken data is taken into account.

When it comes to productivity, the phenomenon of morphological productivity has long been discussed from many perspectives in the literature (see, e.g., Schultink (1961), Bauer 1983, Plag (1999). Yet it has long remained as "one of the central mysteries of derivational morphology" (Aronoff 1976: 35). The most accepted and cited definition of productivity is made by Schultink (1961): Productivity as a morphological phenomenon is the possibility for language users to coin, unintentionally, an in-principle uncountable number of new words by means of the morphological process that underlies the form-meaning correspondence in some words they know. Thus the three main features of productivity are unintentionality, unlimitedness and regularity. The productivity of a language is supported by the frequent use of some regular affixes. The relationship between frequency and productivity has been studied by some scholars (e.g. Baayen 1993; Baayen and Lieber 1991), who proposed quantitative measures of productivity defined with respect to the frequency of a given word-formation process. In other words, frequency is claimed to be a relevant parameter in order to evaluate the productivity and the availability of word formation processes (Gaeta & Ricca 2003).

In this paper, the frequency of derivational and inflectional suffixes in child literature will be determined by conducting a frequency analysis and some conclusions in terms of the productivity of suffixes in children's acquisition of Turkish language will be reached.

1.2 Derivational and Inflectional Suffixes

Turkish, which is an agglutinative language belonging to the Ural-Altaic language family, consists of morphologically fixed word stems and productive suffixes attached to the end of them. The suffixes which are arbitrarily attached to roots or stems of a noun or verb to derive new words and create new meanings are called derivational suffixes. Suffixes that establish temporary semantic relationships within and between words are called inflectional suffixes (Korkmaz, 2009). Derivational and inflectional suffixes are the cornerstones of the morphological structure of Turkish. These suffixes and their sub-categories are shown in Table 1.

Table 1: Derivational and inflectional suffixes of Turkish

Inflectional Suffixes		Derivational Suffixes
Nominal Inflectional Suffixes	Verbal Inflectional Suffixes	1. Suffixes that attach to nominals to form nominals
1. Case Suffixes	1.Tense/Aspect/Modality Suffixes	2. Suffixes that attach to nominals to form verbs
2. Possessive Suffixes	a) Tense Suffixes	3. Suffixes that attach to verbs to form nominals
3. Genitive Case Suffix	b) Modality Suffixes	4. Suffixes that attach to verbs to form verbs
4. Plural Suffix	2. Personal Suffixes	
5. Interrogative Particle	3. Negative marker	
6. Relative Pronoun -ki	4. Interrogative Particle	
7. Copular Markers	5. Copular Markers	

In this study, examples from children's literature were analyzed and a corpus was formed by using the derivation and inflectional suffix patterns in the table.

1.3 Corpus and Corpus Studies in Turkish

The collection of large amounts of spoken or written natural language patterns in a language is called corpus. This corpus created is analyzed quantitatively and qualitatively and analyzed syntactically, lexically and morphologically within the scope of corpus linguistics (Biber, 2004). As a result, the frequency of the structure examined, that is, the number of times that linguistic element occurs is checked.

When we look at corpus studies in Turkish, one of the most basic studies is the frequency dictionary of Turkish written by Göz (2003). Spoken Turkish Corpus (STC) demo version, which has been constructed by the Middle East Technical University, is a spoken corpus study consisting of Turkish conversations (see http://std.metu.edu.tr/). The aim of this Project is to construct a linguistically analyzed corpus consisting of one million words of face-to-face or mediated interactions in contemporary Turkish.

In addition to comprehensive oral and written natural language corpus studies, there are also corpus studies examining inflectional and derivational suffixes, which are the subject of this study. Aksan and Mersinli (2010) examined the corpus that is being formed within the scope of the Turkish National Corpus Project (TNC) in terms of inflectional and derivational suffixes. In addition, Uzun et al. (1992) made a very comprehensive corpus study on derivational suffixes in Turkish. If we look at the corpus of special fields for language acquisition in children, one of the most remarkable studies is the Turkish Child Literature Corpus (TCLC), which is currently being created by Mersin University (see http://turkcederlem.mersin.edu.tr/cocuk/). The aim of this project is to determine the primary expression existence; to analyze the levels of age and readability of child literature works; to identify the lexical areas of child literature sample, and to classify the properties of morphemic, lexeme and syntactic categories. Apart from this ongoing project, no study related to corpus of early childhood literature and its morphological features has been found.

2. Method

This study is a holistic study in which the quantitative frequency analysis was conducted. The data obtained from children's books were analyzed in terms of their morphological characteristics. In addition, derivational suffixes, inflectional suffixes and their subcategories were examined. Finally, the frequency values of these suffixes were determined and a comparison between the categories was made.

2.1 Research Data

All affixes in the 5 books examined in this study were identified and a total of 1503 affixes were examined. Afterwards, these affixes were divided into derivational and inflectional suffixes and their sub-categories and

analyzed together. As a result, 1155 inflectional suffixes and 348 derivational suffixes were examined in detail and the frequency values of these groups were calculated.

2.2 Data Collection

The data of this study were compiled from 5 randomly selected children's books from different publishers. The corpus includes the collection of 3 authors by 3 publishers (Çiçek Publishing, Altın Kitaplar and Yapı Kredi Publications). There were three age groups for the books: 0 to 3 year-olds, 2 to 6 year-olds and 3 to 8 year-olds. The target audience was determined by the publishers, so no additional classification has been done in this study. Selecting the books that make up the data from different publications and age groups is to reach the preschool children's literature more broadly and to diversify the sample.

In this study, it is assumed that children's books would be a sample of child language. We predict that suffixes with high frequency would be acquired early and those suffixes which have low frequency values would be acquired late.

3. Findings and Discussion

The suffixes analyzed in the corpus formation study of the early childhood period Turkish literature were integrated into two categories, namely inflectional and derivational suffixes. Accordingly, the suffixes and their sub-categories are shown in Table 2.

Table 2: Frequencies of use of inflectional and derivational suffixes and their subcategories in early childhood books (n=1503)

Inflectional Suffixes (1155)				Derivational Suffixes	f
Nominal Inflectional Suffixes	f	Verbal Inflectional Suffixes	f	1. Suffixes that attach to nominals to form nominals	41
1. Case Suffixes	345	1. Tense/Aspect/Modality Suffixes	253	2. Suffixes that attach to nominals to form verbs	50
2. Possessive Suffixes	166				
3. Genitive Case Suffix	74	a) Tense Suffixes	214		
4. Plural Suffix	92	b) Modality Suffixes	39	3. Suffixes that attach to verbs to form nominals	173
5. Interrogative Particle	7	2. Personal Suffixes	121		
6. Relative Pronoun –ki	2	3. Negative marker	37		
7. Copular Markers	34	4. Interrogative Particle	4	4. Suffixes that attach to verbs to form verbs	84
		5. Copular Markers	20		
Total	720		435		348

When we look at the general results, a total of 1503 suffixes have been reached in the corpus study of which 1155 are inflectional and 348 are derivational suffixes. It is seen that the number of inflectional suffixes in the corpus is significantly higher than the derivational suffixes. When we look at the inflectional suffixes, nominal inflectional suffixes are used more than the verbal inflectional suffixes, with a frequency of 720. While the number of use of case suffixes is significantly higher (345) among the nominal inflectional suffixes, the most frequently used verbal inflectional suffixes are tense suffixes (214) which are under the category of tense, aspect and modality. The analysis of derivational suffixes showed that suffixes that attach to verbs to form nominals (173) are used most frequently. In general, the derivational suffixes attached to verbs are used more than those that are attached to nominals. According to the data obtained from the corpus and summarized in Table 2, the following results can be obtained:

(a) In early childhood books inflectional suffixes are given more weight than derivational suffixes. This may be because inflectional attachments are more basic, frequently used, more productive, and acquired earlier.

(b) The fact that the usage of derivational suffixes is less in children's literature may result in that these suffixes are less productive and acquired later in the language acquisition process of children.

(c) According to the data obtained from the early childhood literature, the most frequently used and most productive suffixes are the case suffixes. This is followed by tense suffixes, possessive suffixes and personal suffixes, respectively. This information can give us a clue about the basic morphemes of words and sentences formed by children during the acquisition of Turkish.

3.1 Inflectional Suffixes

As a result of the analyzed data, inflectional suffixes were determined in the children's literature; it has been divided into two categories as nominal and verbal inflectional suffixes and examined according to their subcategories. Attachments and their frequency of use are shown in Table 3 and Table 4.

Table 3: Frequencies of nominal inflectional suffixes

Plural Suffix	f	Case Suffix	f	Possessive Suffix	f	Genitive Case Suffix	f	Interrogative Particle	f	Relative Pronoun	f	Copular Markers	f
-LER (92)	92	Dative -E (146)	146	1st person singular -M	5	-(N)IN	74	MI	7	-KI	2	-(I)DI	14
		Accusative -I	71	2nd person singular-(I)N	34							-(I)MIŞ	0
		Locative -DE	57	3rd person singular -I	97							-(I)SE	1
		Ablative -DEN	38	1st person plural -MIZ	6							-DIR	19
		Instrumental -LE	33	2nd person plural -NIZ	4								
				3rd person plural -LERI	20								

Table 3 gives the frequency of inflectional suffixes attached to nominals. Accordingly, -LER plural suffix, case suffixes, possessive suffixes and -(N)IN genitive case suffixes are suffixes with high frequency of use in the corpus. -MI interrogative particle, -KI relative pronoun and copular markers that give predicate to nouns are used less frequently. It is seen that the frequency of use of the -E suffix, which is the marker of dative case among the case suffixes, is significantly higher. The use of case suffixes following the dative suffix is accusative, locative, ablative, instrumental case suffixes respectively.

From this finding, it can be concluded that the dative suffix is highly productive and is important and prioritized for the input that children receive in the lexical acquisition process. When the possessive suffixes are examined, the third person singular possessive suffix -I is present with a significant difference, while the usage level of the second singular -(I)N and third-person plural -LERI possessive suffixes is moderate. The frequency of use of other personal suffixes is very low. From this finding, it can be concluded that third-person narration is mostly used in the analyzed children's books. Frequent reference to possessive and genitive suffixes can provide information on the use of noun phrases in the reviewed literature. When the copular markers attached to nominals are examined, we see that the -DIR and -DIK suffixes are used. The fact that the frequency of the copula used in the construction of the nominal sentences is low may indicate that the nominal sentences are few and the verbal sentences are more in the analyzed books.

As a result, it is observed that the most frequently used nominal inflectional suffixes are dative case suffix -E, third person singular possessive suffix -I, plural suffix -LER, and accusative case suffix -I.

Table 4: Frequencies of verbal inflectional suffixes

Tense Suffixes	f	Modality Suffixes	f	Personal Suffixes	f	Negative marker	f	Interrogative Particle	f	Copular Markers	f
Aorist tense -(I)R	32	Necessity -MELI	10	1st person singular - M	11	-MA	37	MI	7	Past copula -(I)DI (16)	16
Present tense-YOR	30	Desiderative -SA	0	2nd person singular -N / SIN	26					Evidential copula -(I)MIŞ	0
Past tense -DI	140	Optative -A	11	1st person plural - IZ / K	22					Conditional copula -(I)SE	4
Evidential Past tense - MIŞ	5	Imperative (No affix)	18	1st person plural - LIM	4						
Future tense - ECEK	7			2nd person plural - SINIZ / IN	17						
				3rd person plural - LER	41						

Table 4 shows the analysis and frequency of verbal inflectional suffixes in the corpus. According to this, among the inflectional suffixes attached to verbs, the most frequently used ones are tense suffixes and personal suffixes. Other suffixes with less frequency are modality suffixes, -MA negation marker, copular markers and MI interrogative particle.

The past tense suffix –DI, which is among the tense suffixes in the corpus, has a very high frequency of use (140). This finding indicates that the form of expression in the analyzed books is the past tense. This indicates that the productivity level of the suffix is very high, which may prioritize the child acquisition of the –DI suffix. Other tense suffixes used in the corpus are the aorist suffix -(I)R and the present tense suffix -YOR. The future tense – ECEK and evidential past tense suffix –MIŞ are used very little.

Among modality suffixes, the imperative mood which has no suffix is used the most followed by -A optative mood and -MELI imperative mood, respectively. The fact that the modality suffixes are not used much in the corpus and the imperative mood is more may be due to the simplicity of the syntactic structure of the written language which has been studied.

When the personal suffixes in Table 4 are examined, the third person plural suffixes are mostly seen. Since there is no third person singular suffix, it is not included in the corpus. Since third-person narration is dominant in the analyzed data, it is usual not to use other personal suffixes.

The most commonly used copular marker is the –DI suffix. The fact that the frequency of copula is not high may mean that copular structures transform verbs into compound verbs, and that such complex structures are not used much in early childhood books and are produced later in the acquisition process.

As a result, the most commonly used verbal inflectional suffixes are: the past tense suffix -DI, the negative suffix -MA, the 3rd person plural suffix -LER, the aorist suffix -(I)R, and the present tense suffix -YOR.

3.2 Derivational Suffixes

In addition to inflectional suffixes, derivational suffixes were also compiled in the early childhood books examined in this corpus study. The examined derivational suffixes, their subcategories and frequency of use are shown in Table 5 in detail.

Table 5: The corpus results of derivational suffixes and their subcategories in early childhood books

Suffixes that attach to nominals to form nominals (n=41)	f	Suffixes that attach to nominals to form verbs (n=50)	f	Suffixes that attach to verbs to form nominals (n=173)	f	Suffixes that attach to verbs to form verbs (n=84)	f
-LI	18	-LE	30	-MA	49	-(I)R	20
-LIK	10	-LEN	8	-DIK	29	-T	17
-CI	5	-LEŞ	6	-MAK	15	-L	13
-SIZ	3	-R	2	-IP	15	-N	13
-CIK	2	-T	2	-DIĞINDA	9	-DIR	8
-KI	1	-K	1	-KEN	7	-Ş	7
-TI	1	-A	1	-NTI	6	-ALA	4
-CI	1			-A	6	-I	2
				-INCA	5		
				-EREK	4		
				-GE	4		
				-IŞ	4		
				-AN	4		
				-ECEK	3		
				-MIŞ	3		
				-K	2		
				-I	2		
				-CE	1		
				-GIN	1		
				-N	1		
				-DIKCA	1		

When we examine the results given in Table 5, it is seen that the frequency of use of suffixes forming nominal from verb is very high (173), followed by suffixes forming verb from verb (84). The frequency of use of derivational suffixes which nominalize or verbalize the attached noun is low. Among the suffixes that attach to nominals to form nominals, the most productive ones are -LI AND -LIK. For example, kar-lı (snow-y), renk-li (color-ful), dost-luk (friend-ship), kaya-lık (rock-y). Among the suffixes that attach to nominals to form verbs, the suffix -LE is used significantly such as su-la (irrigate), kilit-le (lock). Among the suffixes that form a nominal from the verb, the most frequently used suffix is -MA. Besides the use of forming a noun by changing the meaning of the verb as in the word uçurt-ma (kite), there is also the use as verbal which serves as an infinitive as in the words uç-ma (fly -ing), hatırla-ma (remember –ing). In the corpus created, -DIK, -MAK and -IP suffixes were frequently included in addition to the -MA suffix. –MAK (-to and -ing) suffix mostly functions as verbal and is attached to the verbs as infinitive. Similarly, -DIK suffix is attached to verbs and derive adjectives from them, and -IP suffix derives adverbs from verbs. For example, otur-mak (to sit), koş-up geldi (he came running), kullan-dık-ı kalem (the pen which he use). Among the suffixes that form a verb from verb, -(I)R and -T derivational suffixes are used the most. These two suffixes make the word to which it is attached transitive and causative. For instance, the intransitive verb uç (fly) becomes transitive after getting (I)R suffix as in 'O uçurtmayı uç-ur-du'(He flew a kite). Moreover, as in the sentence 'Dün saçını kestir-t-ti' (She had her hair cut), -T suffix allows to form the causative structure. In addition to these suffixes, –L and –N suffixes which turn the structure into passive and reflexive are among the frequently used suffixes, although less frequently.

4. Conclusion

In this study, randomly selected early childhood literature in Turkish was examined morphologically, derivational and inflectional suffixes were analyzed by examining its lexical features and a corpus was formed. The obtained data were grouped according to derivational, inflectional suffixes and their subcategories, and the frequency of use was found. According to the results, the frequency of use of inflectional suffixes in the books written for children in early childhood period is higher than the derivational suffixes. The dative case suffix –E and the past tense suffix –DI are the most frequently used suffixes and more than other suffixes. Nominal inflectional suffixes are more than verbal inflectional suffixes. In addition, verbal sentences which are performative are included more than nominal sentences. When the derivational suffixes are examined, it has been concluded that the suffixes that are attached to verbs to form nominals are used more, and the suffixes attached to nominals are less. –MA, -LE, and –DIK are the most commonly used derivational suffixes. We can conclude that these structures, which are used more frequently, can be more productive and therefore may have priority in children's vocabulary.

All in all, the analysis of inflectional and derivational suffixes obtained through this corpus study is believed to have helped determine the more frequently used, more productive and prioritized suffixes in the input that children hear and obtain in language acquisition and especially in the process of lexical acquisition. Besides, it is hoped to be helpful by providing a base for further studies.

References

Aksan, A., & Mersinli, Ü. (2010). A corpus based nooj module for Turkish. *The Nooj 2010 International Conference.* Mersin.

Aronoff, M. (1976). *Word formation in generative grammar*. MIT Press.

Baayen, H., & Lieber, R. (1991). Productivity and English derivation: a corpus-based study. *Linguistics,* 29, 801–843.

Baayen, H. (1993). On frequency, transparency and productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology* 1992 (pp. 181–208). Kluwer.

Bauer, L. (2001). *Morphological productivity*. Cambridge University Press.

Biber, D. (2004). *Corpus linguistics: investigating language structure and use.* Cambridge University Press.

Clark, E. (2009). *First language acquisition.* Cambridge University Press.

Gaeta, L., & Ricca, D. (2003). Frequency and productivity in Italian derivation: A comparison between corpus-based and lexico-graphical data. *The Italian Journal of Linguistics, 15*, 63.

Göz, I. (2003). *Yazılı Türkçenin kelime sıklığı sözlüğü.[A frequency dictionary of written Turkish words].* TDK.

Korkmaz, Z. (2009). *Türkiye Türkçesi grameri (Şekil bilgisi). [Turkish grammar (morphology)]*.TDK

Macwhinney, B. (2003). First language acquisition. (Eds.) Aronoff, M., Miller, J. (2003). *The handbook of linguistics.* J. Blackwell Publishing.

Plag, I. (1999). *Morphological productivity. Structural constraints in English derivation.* Mouton de Gruyter.

Schultink, H. (1961). Produktiviteit als morfologisch fenomeen. *Forum der Letteren* 2, 110-125.

Spoken Turkish Corpus (2021, October 16). Middle East Technical University. http://std.metu.edu.tr/

Turkish Child Literature Corpus (2021, October 10).Mersin Üniversitesi. http://turkcederlem.mersin. edu.tr/cocuk/

Uzun, E., Uzun, L., Aksan, M., & Aksan, Y. (1992). *Türkiye Türkçesinin türetim ekleri: Bir döküm denemesi. [Derivational suffixes of Turkish].* Şirin.