A Developmental Study of Turkish Connectives*

Enis G. Oğuz[1], Duygu Özge[2]

ORCID ID: [1]0000-0001-5819-4926, [2]0000-0002-1698-5479

[1]İstanbul Üniversitesi Yabancı Diller Yüksekokulu 34116 Fatih, İstanbul
[2]ODTÜ Yabancı Diller Eğitimi Bölümü 06800, Ankara

[1]enisgoguz@gmail.com, [2]duyguozge@gmail.com

(Received 04 March 2020; accepted 18 June 2020; revised 28 December 2020)

ABSTRACT: Connectives present an interesting phenomenon; while even 3-year-olds can produce them accurately, the real mastery of understanding their meanings takes much more time in language acquisition. Despite its allure, studies focusing on connective acquisition are limited in number, especially in Turkish. In order to investigate the developmental path of different Turkish connectives, we tested the comprehension of temporal, causal, and adversative connectives in 184 Turkish-speaking primary-school children and 20 Turkish-speaking adults. Our results demonstrated that high-achieving 5th graders show an adult-like performance. Regarding connective types, temporal connectives posed a greater challenge for children compared to causal and adversative connectives. The questions were more challenging for all groups (but especially for 3rd graders and low-achieving 5th graders) when they require a backward temporal connective (before that) as the correct answer or include object-referred clauses. We explain these results in light of the multi-dimensional account of connective acquisition.

Key words: Turkish, connective acquisition, connective types

Türkçe Bağlaçlar Üzerine Gelişimsel Bir Çalışma

ÖZ: Çocuklar üç yaşındayken bile bağlaçları doğru olarak kullanabilirken, anlamlarını tam manasıyla edinmeleri ve anlamlandırmaları çok daha uzun zaman gerektirmektedir; bu bakımdan, bağlaçlar gelişimsel açıdan ilginç bir olgu sunar. Cazibesine rağmen, bağlaç edinimi üzerine yoğunlaşan çalışmalar sayı olarak oldukça sınırlıdır -özellikle de Türkçede. Bu çalışma Türkçe bağlaç edinimindeki gelişimsel rotayı incelemek amacıyla zamansal, nedensel ve karşıtsal bağlaçların anlamlandırılma süreçlerini anadili Türkçe olan 184

* This is a amended version of the following study: Oğuz, E. and Özge, D. (2020). A developmental study of Turkish connectives. *Dilbilim Araştırmaları 31*(1), 77-100. http://dx.doi.org/10.18492/dad.698824.

http://dx.doi.org/

Dilbilim Araştırmaları Dergisi, 2020/2, 339-361.

ilkokul çağında çocuk ve 20 yetişkin ile test etti. Sonuçlar, akademik başarıları öğretmenleri tarafından yüksek olarak değerlendirilen 5. sınıf öğrencilerinin yetişkin seviyesinde performans gösterdiklerini ortaya koymaktadır. Bağlaç türleri ile ilgili bulgular ise zamansal bağlaçların çocuk grupları için nedensel ve karşıtsal bağlaçlara göre daha zor olduğunu göstermektedir. Doğru cevabı geriye dönük zamansal bağlaç (öncesinde) olan veya nesnelere atıfta bulunan tümce içeren sorular ise tüm gruplar için fazladan zorlayıcıydı (özellikle 3. sınıf öğrencileri ve akademik başarıları düşük olarak değerlendirilen 5. sınıf öğrencileri için). Elde edilen bu bulgular bağlaç ediniminde çok boyutlu anlamlandırma bakış açısını destekler niteliktedir.

Anahtar sözcükler: Türkçe, bağlaç edinimi, bağlaç türleri

1 Introduction

"John has many pets. He has a dog. He has two cats. He has a parrot. He has a hamster. He has so many animals. He loves them. He does not want to be alone at home. He organizes his life around these animals. He wakes up very early to feed them. He feeds the dog. He feeds the cats. He feeds the parrot. He feeds the hamster. Not many people can afford the extra energy these animals require. John loves it. It became his lifestyle."

Despite this simple paragraph makes all the sense, it does not flow because the ideas are largely disconnected. The same text becomes much easier to follow when the utterances are linked with connective devices as follows:

"John has many pets. He has a dog, two cats, and a parrot. He also has a hamster. He has so many animals because he loves them, and he does not want to be alone at home. He organizes his life around these animals. For instance, he wakes up very early to feed them. He feeds the dog and the cat before he feeds the parrot and the hamster. Not many people can afford the extra energy these animals require but John loves it because it almost became his lifestyle."

It is almost like a reflex for adults to produce connected speech, and children surprisingly begin to use certain connectives as early as three years of age. However, understanding when exactly they attain the full mastery of their meaning and function is crucial because this ability may be linked to later expressive abilities and academic success (e.g. Im-Bolter, Yaghoub Zadeh, & Ling 2013). This study focuses on the comprehension of three types of connectives (i.e., temporal, causal, and adversative) in Turkish-speaking primary-school children between the ages of 8 and 10 as compared to adults. We aim to document (i) how children at this age group perform in their interpretation of connectives as compared to adults, (ii) whether there is any development between age 8 and 10, and (iv) whether there is any difference in the complexity posed by the type of the connective. Below, we summarize why connectives are necessary in a discourse, review previous studies on the

acquisition of connectives and crosslinguistic findings regarding connectives, and then we present our study and results.

1.1 Why are Connectives Necessary?

Connectives, as the name suggests, connect clauses, phrases, and words. They are divided into four broad categories as additive, temporal, causal, and adversative (Table 1). Some connectives can be grammatically removable constituents, such as the connective *and* in (1). The absence of such a connective does not cause a deficiency in grammar or comprehension:

(1) I have a little dog and I like it.
(2) I have a little dog. I like it.

Table 1. Relation Types between Sentences Marked by Connectives (based on Cain and Nash, 2011)

Category	Example	Explanation
Additive	He has a jet-plane on his own. *Also*, he possesses three sports cars.	Additive relations indicate that each new sentence is adding similar information to what is already presented so far.
Temporal	Brush your teeth *before* you go to bed.	Temporal relations present temporal relations between events.
Causal	I failed the test *because* I had not studied a lot.	Causal relations implicate that one of the presented segments is the reason for the other segment, which is the consequence.
Adversative	Jack hates spinach; *however*, he eats it as he is on a diet.	Adversative relations are perceived as two segments having a contrastive relationship.

They can also be irreplaceable pieces of an utterance, like the *or* in the following sentence (3). Removing such connectives will result in ungrammaticality, such as in sentence (4):

(3) Tom can't read or write.
(4) *Tom can't read write.
(5) Tom can't read. He can't write.

Some may argue that the example (3) is not a solid argument for the irreplaceability of some connectives since creating an independent sentence with the second verb *write* would result in a grammatical utterance (5). Although the sentence (5) is grammatical, the sentence (3) is easier to process. The removal of

the connective *or* might be compensated by inserting another connective *and* for this sentence; however, the meaning would not be the same.

(6) Crows are usually black. This one is white.
(7) Crows are usually black but this one is white.

Similarly, the connective *but* in the sentence (7) does not influence the grammaticality of the sentence, yet it signals an upcoming contradiction. These examples show how connectives can improve the coherence of a text. Connectives, then, contribute to the relationship between sentences, grammaticality, ideas represented, as well as giving signals about what is coming next. By doing so, they facilitate processing and comprehension (e.g. Maury & Teisserenc, 2005; Sanders, Land, & Mulder, 2007; van Silfhout, Evers-Vermeul, & Sanders, 2014).

Previous studies state that people have a cognitive bias towards continuity and causality (e.g. Murray, 1997). For example, the causality-by-default hypothesis argues that when people read texts, they try to create the most informative presentation by initially assuming two sequential sentences are in a causal relation (for a detailed review, see Sanders, 2005). The findings showing that explicit markers are much more common with discontinuous or temporal discourse relations (Murray, 1997; Asr & Demberg, 2012) further support the causality-by-default hypothesis. In other words, people expect continuity while reading a text, and the violation of this continuity requires explicit marking. If the first clause of a sentence refers to a more recent time than the subsequent clause, for example, this violates the continuity of the text and necessitates an explicit indicator, as in (8). Differences do not only arise with the change in the connective types; the polarity within the same connective category also influences the rate of explicit marking. Polarity refers to the nature of a causal relationship. While a positive polarity refers to a forward temporal relationship (9), a negative polarity refers to a backward temporal relationship (8). For instance, Asr & Demberg (2012) showed that temporal orders with positive polarity are more often left implicit as they are easily inferred by people compared to temporal orders with negative polarity.

(8) Before she had breakfast, Sara washed her hands.
(9) After he watched the show, Kevin returned home.

Considering these, we can say that adults easily integrate the meaning contributed by a connective in an utterance. Even in the absence of a connective, the mind assumes some default meaning relations trying to fill in the meaning and the coherence gap. This is crucial not only in conversation as it would ensure successful communication but also in reading as it would ease the comprehension

of a text. Even if connective acquisition is an important part of language acquisition, we have less information about how this ability develops, when children reach an adultlike interpretation of connectives, and how different connectives influence comprehension.

1.2 How Do Children Acquire Connectives?

Even 3-year-old children produce connectives successfully. Blything, Davies, and Cain (2015) examined the comprehension of temporal connectives in 3- to 7-year-old English-speaking children and whether polarity had any effect on their comprehension. They used a connective comprehension task, where each child sees two animations referring to different actions for each trial. These two animations freeze on the screen, each occupying the half size of the screen. After that, the child listens to a sentence including two clauses, each reflecting one of these actions along with a connective tying those clauses together. The actions have a sequential relationship: One is done after the other. Depending on the connective, the first action is told as either the first clause or the second clause. The child is asked to point at the animation showing the first action for the given sentence. The children aged between 3 and 4 showed an above-chance performance, while the children aged between 6 and 7 showed a highly accurate performance.

Despite the high performance of the older children in this study, there are other studies showing that achieving the complete mastery of connectives requires much more time. For instance, Cain, Patson, and Andrews (2005) found that English-speaking children aged between 7 and 10 were far from adult-like performance. In a later study, Cain and Nash (2011) also argued that although 10-year-olds in their study could perform at adult performance in some tasks, their performance was still behind adults in others. Thus, they deduced that connective skills were still developing despite children could produce them accurately beginning from early ages.

Two different accounts try to explain the road to connective mastery: Bloom's cumulative complexity account and Evers-Vermeul and Sander's multi-dimensional account. The cumulative complexity account goes back to the 1980s; Bloom, Lahey, Hood, Lifter, and Fiess (1980) made an influential connective acquisition study with four children (around 24 months old) for over 10 weeks, according to which simpler semantic relations were acquired earlier than more complex ones. More specifically, children acquired additive relationships first, which was followed by the acquisition of temporal relationships. The latest acquired relationships were causal and adversative relationships.

The multi-dimensional account partly supports this view. Investigating the emergence of connectives in Dutch via CHILDES corpus (MacWhinney,

2000), Evers-Vermeul and Sanders (2009) rejected a one-dimensional explanation of semantic relations (additive < temporal < causal < adversative). Their investigation suggested that conceptually less complex connectives are acquired before relatively more complex ones. As opposed to the connectives showing only continuity (i.e. indicating additive relationship for upcoming sentences) (e.g. *and*), the connectives indicating a multiple set of relations such as continuity, temporality, and causality at once (e.g. *consequently*) were considered as conceptually more complex. Evers-Vermeul and Sanders (2009) further advocated that even the nature of a specific relation was determinant in conceptual difficulty; connectives with negative polarity, for example, were considered as more complex than connectives with positive polarity. Thus, while the cumulative complexity account suggests that connective acquisition follows a straightforward order based on semantic relationship, the multi-dimensional account proposes that multiple factors such as semantic relationship and polarity collectively determines the order of connective acquisition.

Although having different claims, both theories agree that different connectives are acquired at different times during the course of language acquisition. Even after the acquisition, the processing difficulty caused by each connective type might still differ. The next section summarizes the findings related to this phenomenon.

1.3 Are All Connectives Produced with Same Ease?

Several studies have found processing differences for different types of connectives. Adversatives, for example, have been reported to be more demanding to interpret even for adults (e.g. Goldman and Murray, 1992; Murray, 1997; Köhne and Demberg, 2013). As one of the early studies, Murray (1997) found that sentences with adversative connectives were rated as less coherent compared to sentences with additive and causal connectives. Also, sentences with adversative connectives required longer times for comprehension. In a more recent study, Köhne and Demberg (2013) conducted two experiments to examine the processing of adverbial and causal connectives. The results of the first experiment suggested that processing sentences with adversatives led to higher cognitive processing demands and less accuracy in comprehension. The second experiment, however, failed to present significant differences between adversative and causal connectives. Köhne and Demberg (2013) argued that casual connectives might lead to rapid predictions if their context is clear and constraining. As for adversatives, a relatively larger argument scope could explain the findings. What is meant by a relatively larger argument scope is the following: Consider the utterances in (10) and (11); while the causal connective in (10) is most likely to continue with A, the adversative connective in (11) have

similar chances of continuing with either A or B, which is offering a less precise -thus larger- argument scope. These two studies seem to support the cumulative complexity account; since processing a connective with a larger argument scope is more challenging even for adults, this and other factors are likely to also affect connective acquisition.

(10) (Causal) Jack wants to go to A and B. A is more beautiful. Therefore…
(11) (Adversative) Jack wants to go to A and B. A is more beautiful. However...

It is important to underline that the possibility of a relatively larger argument scope does not always necessarily mean more ambiguity for upcoming information. Murray (1995) advocated that adversative connectives are also highly constraining for what comes next; the text after an adversative connective is likely to contrast with the text before the connective. This highly-constraining nature might especially be helpful for the fill in the blank questions or multiple-choice questions asking for the appropriate connectives -since the arguments for second clauses are already given. Support for this argument comes from the study of Cain et al (2005); the children aged between 8 and 10 scored significantly higher for additives and adversatives compared to casuals and temporals in cloze tests. The children first listened to two stories in which some connectives were taken out. The tests required children to select appropriate connectives among the options. The sentences were taken from the stories they had listened to. It is likely that children could have recognized the need for adversatives better when they were absent, as two contrasting clauses usually require explicit marking. The finding is also interesting in that it contrasts with the aforementioned studies showing extra costs for adversative connectives (Köhne and Demberg 2013; Murray, 1997). Therefore, we can conclude that there is no consensus on whether adversative connectives are more difficult to process compared to causal connectives.

As for temporal connectives, Crain, Shankweiler, Macaruso, Bar-Shalom (1990) found that children with poor reading skills had problems with negative polarity possibly due to high memory demands; poor readers made more mistakes with the sentences like (12) than the ones like (13).

(12) Push the red car before you push the largest horse.
(13) Push the smallest horse after you push the blue car.

Memory load and memory skills are potential factors in processing connectives. Crain et al. (1990) argued that their findings could not be due to limited linguistic knowledge since the same participants showed an overall good performance with temporal connectives. Considering the high error rates for the sentences with negative polarity, Crain et al. (1990) inferred that heavy load on working memory

was the reason behind the results. Blything et al. (2015) used a digit-span task to measure memory skills and similarly found that clause order and memory skills were essential in the comprehension of temporal connectives in a study with 3- to 7-year-old children. Participants were more accurate with positive polarity while working memory skill was indicative of their performances. The pattern from Crain et al. (1990) and Blything et al. (2015) is in line with the pattern in Evers-Vermeul and Sanders's (2009) study, which showed that negative polarity connectives are acquired later than positive polarity connectives.

As it has been shown, not only the meaning of the connective but also its polarity might be an important factor in its acquisition and processing. This is especially important for the present study as the lower performance for some connectives might also stem from the cognitive load and memory limitations.

Another potential factor influencing connective processing might be cross-linguistic factors: Differences in syntax, morphology, or phonology might also lead to differences in acquiring and processing certain connectives in different languages. This phenomenon is discussed in the next section.

1.4 Crosslinguistic Factors in Connective Acquisition

Previous studies indicate that there are cross-linguistic factors contributing to connective use and connective acquisition. Comparing the use of English and Japanese causal markers in conversations, Ford and Mori (1994) found that while English had a single causal marker (because) to account for disagreement, Japanese had two causal connectives for the same purpose (*kara* and *datte*). While the connective *datte* indicates a more severe disagreement, the connective *kara* is perceived as an invitation for negotiation and collaboration to overcome the disagreement. Therefore, Japanese allows the expression of different types of disagreement via two different connectives. Altenberg (2007) examined Swedish translations of English texts in the English-Swedish Parallel Corpus (ESPC). The corpus had the original texts in both English and Swedish along with their translations for the other language. While the frequency for explicit causative marking was higher for original Swedish texts compared to original English texts, the translations from English to Swedish showed a tendency for *zero translation* for causal connectives, meaning that causal relations were indicated more implicitly in the English versions. The findings were attributed to cross-linguistic differences for connective use: Swedish needs more explicit marking for showing causative relations, while English uses fewer connectives to show the same type of relations.

As for acquisition studies, research by Evers-Vermeul and Sanders (2009) compared their findings based on a Dutch corpus with the findings from Bloom et al. (1980) that was based on an English corpus. For a connective to be counted as acquired, the criteria required the use of that connective to be

grammatical and creative (i.e. not appearing as a part of a fixed expression). The comparison revealed that Dutch children followed different acquisition routes for certain Dutch connectives (en, maar, toen, and want) than English children. This diversity was attributed to differing conceptual difficulty between connective equivalents in English and Dutch. As a contradictory result against the cross-linguistic effects in connective acquisition, Zufferey, Mak, and Sanders (2015) compared 5-, 6-, 7-, 8- year old Dutch and French children for their comprehension of subjective relations (relations between claims and conclusions) and objective relations (relations between causes and consequences) in stories. The Dutch language generally used a different causal connective to tie objective relations while modern French used the same connective for both object-oriented and subject-oriented causal relations between sentences. Although this was taken as a possible negative influence on the performance of French children, the results showed that French and Dutch children did not differ in their scores. Overall, it seems there is no agreement on whether crosslinguistic differences influence the patterns of connective acquisition.

1.5 Why Do We Need the Present Study?

The present study investigates a developmental phenomenon that has not been well-documented due to the limited number of studies on the acquisition of connectives even in well-studied languages. As reviewed above, we do not have a consensus about the order of acquisition; while Bloom's cumulative complexity account (Bloom et al. 1980) suggest a straightforward acquisition order based on merely semantic relations, Evers-Vermeul and Sanders' (2009) cumulative complexity approach (2009) states that there are more factors in determining the complexity of a connective and its acquisition order. The current study contributes to the literature by offering an insight into the development of connectives from a language that has not been studied extensively before. Aksu (1978) examined the speech of Turkish-speaking children (2;0 to 4;6) and investigated the acquisition stages of causal connectives. Similarly, Sofu and Şimşek (2016) investigated the patterns of Turkish-speaking children's (2;0 to 6;0) use of additive connectives throughout the acquisition of coordination. Although providing crucial insights into the natural development of causal and additive connectives in child speech in Turkish, these studies did not compare different types of connectives. Also, the developmental studies in Turkish so far have focused on the first use of connectives mostly in the production domain; these studies have not investigated when children attain adult-like performance in comprehending different types of connectives, which is the main focus of our study. Studying Turkish-speaking children is crucial from yet another perspective as well. Turkish-speaking children achieve remarkable reading skills

at the end of the first grade as a result of the language's shallow orthography and simple syllable structure (Durgunoğlu & Öney, 1999; Öney & Durgunoğlu, 1997) and acquire causal connectives at early ages due to the transparency of the language (Aksu-Koç & Slobin, 1985). The relatively faster acquisition of connectives and reading skills compared to many other languages might also allow a faster improvement in discursive awareness (e.g. coherence and cohesion), which might, in turn, lead to a faster mastery of connective use. Thus, Turkish-speaking children might also show adult-like connective comprehension skills at earlier ages.

The studies investigating the adult-like mastery of connective use in other languages had several limitations. For instance, Cain and Nash (2011) could not compare the scores of different connectives types since they included *and* connective as an option in the cloze test with multiple choices. *And* can be used instead of other connectives. When *and* appeared with its interchangeable connectives in the options for a question, the question usually had two acceptable answers. In such incidents, the participants preferred choosing *and* connective, which prevented a comparison among connective types due to high reliance on this connective. Another study by Cain and Nash (2005) did not report the reliability scores for the connective tests. Our study aims to fill in these limitations by (i) comparing different connective types, (ii) presenting only two options in the cloze test and not having *and* among the alternative responses, and by (iii) ensuring the reliability of the scores.

Our research questions in this study are as follows:

1. Are there any differences among 3rd graders, low-achieving 5th graders, high-achieving 5th graders, and adults in terms of overall connective scores?
2. Are there any differences among 3rd graders, low-achieving 5th graders, high-achieving 5th graders, and adults in terms of different connective type scores?
3. Are there any differences among different connective type scores within 3rd graders, low-achieving 5th graders, high-achieving 5th graders, and adults?

To address these questions, we designed a cloze-test study modeled after Cain and Nash (2011), where children read two-sentence utterances in which a connective is replaced by a blank and chose the right connective among the alternatives.

2 Method

2.1 Participants

Sixty 3rd graders (age range: 8-9 ; mean age:8.76),124 5th graders (age range: 9-10; mean age: 10.74), and 20 adults (age range: 22-35, mean age: 27.88) participated in the study. Teachers from two state schools were included in the study as adult participants. The initial intention was to include equal numbers of 3rd and 5th graders; however, as the Turkish teachers in the middle-school stated that there were considerable differences among the academic performances of 5th grader classrooms, we created two groups of 5th graders based on the evaluation of the Turkish teachers: high-achieving (HA) and low-achieving (LA). Then, two classes were randomly selected from each group. Sixty-two 5th graders from the high-achieving classes and the same number of students from the low-achieving classes participated in the study. Two 3rd graders, 4 LA 5th graders, and 1 HA 5th grader did not complete the test. In addition, we excluded the outliers having a score above or below 2 standard deviations of the group mean. The significant findings in the analyses did not change with or without the outliers. The final number of the participants was 190 with 55 3rd graders, 56 LA 5th graders, 59 HA 5th graders, and 20 adults. All participants had normal or corrected-to-normal visions, and no of them had any cognitive disorders. They were all monolingual native speakers of Turkish.

2.2 Materials

We conducted a comprehension study modeled after Cain and Nash (2011). A multiple-choice question exam was created in which children were asked to choose the best fit for the blanks in small discourses. Cain and Nash (2011) used a three-option test with a fixed *and* option across connective types, and this did not allow them to compare the performances between connective types since most items had two acceptable answers (*and* and a specific connective). Thus, we created a multiple-choice test with forced-choice items (i.e. multiple-choice items with two options) for this study without including the connective *and* as an option. The test consisted of 48 items in 3 conditions: temporal connectives, causal connectives, and adversative connectives (Table 2).

Two connectives were selected for each connective type (temporal, causal, and adversative). Sentences were created by tying two clauses by a connecter in the middle (Table 2). The first clauses in the sentences included a subject, an inanimate object, and a verb. The second clauses were shorter and did not have any overt pronouns. In half of these sentences, the second clauses referred to the subjects of the first clauses, while in the other half, they referred to the objects of the first clauses. In Turkish, null pronouns are usually linked to subjects (Turan, 1996). Şen (2019) provided empirical data for this pattern by examining Turkish novels; the findings showed that null pronouns are linked more to third-person singular subjects compared to other antecedents in Turkish

texts. This tendency might lead Turkish readers to interpret null pronouns initially as subject anaphors. Although it was not the focus of the study, such a comparison would allow us to test whether the object-referring null pronouns would cause any difficulty in the comprehension of connectives.

Table 2. Example Items from the Multiple-choice Test (s.= suffix, acc.=accusative, poss.=possessive)

Category	Example Questions	Options
Temporal	Tolga kulaklığını çıkardı……müzik (Tolga)(earphones+poss.s.+acc.s.) (take off+past tense s.) (music) dinliyordu. (listen+past continuous tense s.) *Tolga took off his earphones……(he) was listening to music.*	a) bu yüzden (*therefore*) b) öncesinde (*before that*)
Causal	Melek ödevini yapmadı ……. (Melek) (homework+poss. s.) (do+negation s. + past tense s.) dersten kaldı (course+ablative s.) (fail+ past tense s.) *Melek didn't do her homework……..(she) failed the class.*	a) yine de (however) b) bu yüzden (*therefore*)
Adversative	Eren oyuncağı istedi……. satılık değildi. (Eren) (toy+acc. s.) (want) (sale) (was+ negation s.) *Eren wanted the toy……..(it) wasn't for sale.*	a) ama (*but*) b) çünkü (*because*)

An equal number of male and female names were used in all conditions (e.g. connective type, subject/object reference). These names were taken from the most common names in Turkey, according to the Turkish Statistical Institute (2017). Three language teachers studied the sentences and made sure that there were no ambiguities related to pronoun resolution. Each connective appeared 8 times as a correct answer and 12 times as a distractor. A pseudo-randomization software, Mix (van Casteren & Davis, 2006), was used to create two versions of the test such that the same condition did not appear sequentially more than three times, and the same option was not the correct answer more than three times in a row. Also, the sentences in which second clauses refer to the same type of pronoun (subject or object), and the sentences with the same connectives as options were not used more than two times in a row. Finally, two identical tests with different item orders were created from these initial two tests; therefore, we ended up with four different versions of the same test to avoid possible order effects.

First, the test was piloted with 10 adult native speakers to check the naturalness of the utterances. All participants rated the items as natural, so we made no changes. Furthermore, the connectives in the tests were removed, and

all clauses were listed in a paper to check whether 3rd graders could understand the clauses on their own. Thirty 3rd graders checked the clauses; 7 words and 1 name were replaced with simpler items as some of the students were not familiar with them. Such confirmation was crucial since, without it, any result could be due to children's failure to understand clauses rather than their failure to comprehend connectives. A Cronbach's Alpha test was used to check the reliability of the test after data collection. The test was highly reliable for all children groups (3rd graders: 0,865; LA 5th graders: 0,896; HA 5th graders: 0,745).

2.3 Procedure

We got ethical approval from Middle East Technical University. The middle-school was selected for practical reasons as the first author worked there. The school of the 3rd graders was picked randomly. Due to the performance differences between 5th graders, all 5th grade classes in the middle school were divided into two groups as high achievers and low achievers, and then two 5th grade classes were picked randomly from both high achiever and low achiever groups (a total number of 4 classes for 5th grade). The child participants were under the age of 18, so their consent was taken orally. Additionally, their principals and teachers were asked for permission for data collection at school, and their parents signed a consent form allowing their child to participate in our study.

The students were tested in a quiet classroom in small groups. The first page of the test (24 questions) was always given first, which included detailed instructions about the procedure and an example question along with an answer. Then, the students were given 15 minutes for the first session of the task. After a 10-minute break, the students were given the second page (24 questions). This session also took 15 minutes. The order was the same for all participants. For the adults, the test was given in a single session. The same items with the same instructions and the same example were used. A total number of 7 participants (Five 5th graders and two 3rd graders) did not finish the study; their papers were not included in the data analysis. Furthermore, we removed the outliers who performed 2 standard deviations below and above the mean (7 participants). The final number of the participants included in the final data was then 190.

3 Findings

3.1 Data Analysis

The data were analyzed with Jamovi 1.1.9.0. Since the data for some connective type scores were not normal in some participant groups, all the data were transformed† for normality. With two decimals considered, the data for total scores, adversative connectives, causal connectives, and temporal connectives were normal in all child groups based on Skewness and Kurtosis values (George & Mallery, 2010) after the transformation (Adult data was normal for total scores and temporal connective scores). For the data that did not conform to the normality assumptions, non-parametric tests were used.

Due to a printing error in some tests, the answers for a single question (Question 22) were not reliable for some 3rd grader papers (28 papers) and thus removed. The question included the adversative connective *yine de* (however). Missing data rate around 5% or lower is inconsequential (Schafer, 1999); the removed data in our study accounted for 0.003% of the total data. Since the question numbers were equal across connective types, removing all the answers for this question from our dataset was not practical. Excluding the participants would mean losing roughly half of the 3rd graders, which would reduce the statistical power of the study. Instead, we created two 3rd grade groups: missing data group (twenty-eight 3rd graders missing the data for Question 22) and no missing data group (twenty-seven 3rd graders without any missing data). To check the similarity of the test performances between the two groups, we used Mann-Whitney U tests after excluding Question 22. The results showed no significant differences between the two groups in total scores (p=0.219), adversative connective scores (p.=0.227), and the scores for the connective *yine de* (p=0.256). Having decided that the groups were similar, the mean score for this single question was calculated based on the 3rd grader answers without the missing data, and this mean score (0.90) was used for replacing the missing data in the missing data group.

3.2 Total Test Scores

Table 3 shows the descriptive statistics for the connective test. As expected, adults had the highest score. HA 5th graders showed a close performance to adults (mean difference: 0.78). Although 3rd graders performed similar to LA 5th graders (mean difference: 0.11), their mean score differed more from HA 5th graders (mean difference: 3.69) and adults (mean difference: 4.47). Overall, the general pattern revealed similar performances for adults and HA 5th graders, while 3rd graders and LA 5th graders also performed close to each other but worse than the other two groups.

† log10(maximum value+1 – observation)

Table 3. Correct Answer Averages of the Connective Test for each group (Maximum score was 48 points)

Groups	Mean	SD
3rd Graders	43.23	3.79
LA 5th Graders	43.34	4.71
HA 5th Graders	46.92	1.13
Adults	47.70	0.47

A One-Way ANOVA test was used to investigate whether the differences among the groups were significant. The analysis showed significant differences among the scores of the groups [$F(3, 186) = 31.5$, $p<.001$, $\eta^2 = 0.337$]. A Bonferroni Post-hoc test revealed that adults performed significantly better than LA 5th graders [$t(186) = 6.69$, $p<.001$, $d=0.48$] and 3rd graders [$t(186) = 7.18$, $p<.001$, $d=0.52$]; HA 5th graders performed significantly better than LA 5th graders [$t(186) = 6.34$, $p<.001$, $d=0.45$] and 3rd graders [$t(186) = 7.02$, $p<.001$, $d=0.50$]. There were no further significant main effects.

3.3 Different Connective Type Scores

Table 4. Correct Answer Averages for Different Connective Types (Maximum score was 16 for each connective type).

Groups		Mean	SD
3rd Graders	Adversative	15.04	1.17
	Causal	14.69	1.27
	Temporal	13.49	2.33
LA 5th Graders	Adversative	15.23	1.39
	Causal	14.63	1.58
	Temporal	13.48	3.10
HA 5th Graders	Adversative	15.85	0.36
	Causal	15.54	0.62
	Temporal	15.53	0.63
Adults	Adversative	15.95	0.22
	Causal	15.95	0.22
	Temporal	15.80	0.41

To analyze the possible differences between the groups regarding the specific connective types, a Repeated Measures ANOVA was used with repeated measure factors based on connective types (Adversative, causal, temporal) and a between-subject factor with 4 levels based on groups (3rd graders, LA 5th graders,

HA 5th graders, adults). The analysis revealed a significant effect of connective types [F(2, 372)= 28.98, p<.001, η^2=0.048], a significant effect of groups [F(3, 186)=27.01, p<.001, η^2=0.190], and a significant interaction between connective types and groups [F(6, 372)= 3.96, p<.001, η^2=0.020]. There were no further significant main effects or interactions.

Further ANOVAs revealed which graders had significant differences among each other in terms of connective scores. For adversative connectives, a one-way ANOVA analysis showed significant differences among groups [F(3, 186)= 10.93, p<.001, η^2= 0.150]. Further Bonferroni post-hoc tests showed that adults performed better in adversative connectives than LA 5th graders [t(186)= 2.91, p=0.02, d=0.21] and 3rd graders [t(186)=4.14, p<.001]; HA 5th graders performed better in adversative connectives than LA 5th graders [t(186)= 3.22, p=.009, d=0.23] and 3rd graders [t(186)= 4.92, p<.001, d=0.36]. For causal connectives, a one-way ANOVA analysis showed a significant difference among groups [F (3, 186)= 14.05, p<.001, η^2=0.185]. Further Bonferroni post-hoc tests showed adults performed significantly better in causal connectives than LA 5th graders [t(186)= 4.87, p<.001, d=0.35] and 3rd graders [t(186)= 5.10, p<.001, d=0.37]; HA 5th graders performed significantly better than LA 5th graders [t(186)= 3.97, p<.001, d=0.29] and 3rd graders [t(186)= 4.29, p<.001, d=0.31]. Finally, for temporal connectives, a one-way ANOVA analysis showed significant differences among groups [F(3, 186)= 23.60, p<.001, η^2=0.276]. Further Bonferroni post-hoc tests showed that adults performed significantly better in temporal connectives than LA 5th graders [t(186)= 5.30, p<.001, d=0.38] and 3rd graders [t(186)= 5.73, p<.001, d=0.42]; HA 5th graders performed significantly better in temporal connectives than LA 5th graders [t(186)= 5.87, p<.001, d=0.43], and 3rd graders [t(186)= 6.46, p<.001, 0.47]. There were no further significant main effects.

To analyze the differences among the connective types within the groups, paired-samples t-tests were used. 3rd graders scored significantly better with adversative connectives compared to causal connectives [t(54)=2.27, p=.028, d=0.31] and temporal connectives [t(54)=5.98, p<.001, d=0.81]. They also scored significantly better with causal connectives compared to temporal connectives [t(54)=4.01, p<.001, d=0.54]. LA 5th graders scored significantly better with adversative connectives compared to causal connectives [t(55)=3.01, p=.004, d=0.40] and temporal connectives [t(55)=6.03, p<.0001, d=0.81]. They also scored significantly better with causal connectives compared to temporal connectives [t(55)=3.14, p=.003, d=0.42]. HA 5th graders scored significantly better with adversative connectives compared to causal connectives [t(58)=3.78, p<.001, d=0.49] and temporal connectives [t(58)=3.54, p<.001, d=0.46]. For HA 5th graders, no significant difference was found between temporal connectives and causal connectives. And as for adults, no significant differences were found among the connective types.

3.4 Temporal Connectives with Negative and Positive Polarity

Table 5. Correct Answer Averages for the Temporal Connectives across the Groups (Maximum score was 8 for each specific connective).

Groups		Mean	SD
3rd Graders	Before that	6.04	1.72
	After that	7.45	1.02
LA-5th Graders	Before that	6.16	1.96
	After that	7.32	1.39
HA-5th Graders	Before that	7.61	0.62
	After that	7.92	0.28
Adults	Before that	7.80	0.41
	After that	8.00	0.00

As the lowest scored connective type for all groups, temporal connectives *before that* (backward temporal connective with negative polarity) and *after that* (forward temporal connective with positive polarity) were further analyzed. Table 5 shows the differences between the two temporal connectives across the groups. A non-parametric test (Wilcoxon Matched Pairs test) was used to compare the connective *öncesinde* (before that) and *sonrasında* (after that). The results showed significant differences for 3rd graders ($p<.001$, $d=1.03$), LA 5th graders ($p<.001$, $d=0.89$), HA 5th graders ($p=003$, $d=0.42$), and adults ($p=.046$).

3.5 Object-Reference and Subject Reference

Table 6. Scores for the Subject-referred and Object-referred Questions across the Groups (Maximum score was 24 for each reference type).

Groups		Mean	SD
3rd Graders	Subject-Referred	22.33	1.71
	Object-Referred	20.89	2.32
LA 5th Graders	Subject-Referred	22.20	2.65
	Object-Referred	21.14	2.39
HA 5th Graders	Subject-Referred	23.68	0.57
	Object-Referred	23.24	0.93
Adults	Subject-Referred	24.00	0.00
	Object-Referred	23.70	0.47

A final analysis was conducted for the possible differences between the performance for subject-reference and object-reference conditions. Table 6 presents the mean scores for subject-referred questions and object-referred questions across the groups. The results of a Wilcoxon Matched Pairs test showed significant differences for 3rd graders ($p<.001$, $d=0.90$), LA 5th graders ($p<.001$, $d=0.75$), HA 5th graders ($p=.003$, $d=0.42$), and adults ($p=.02$, $d=0.64$).

4 Discussion

We set out to investigate the comprehension of adversative, causal and temporal connectives in primary school children (aged 8-10 years) and adults. We asked children to read two linked utterances with a missing connective and select the correct connective between two given choices. Our results showed that adults and high-achieving (HA) 5th graders had similar scores; they outperformed 3rd graders and low-achieving (LA) 5th graders in overall scores and each connective type score (adversative, causal, and temporal). All child groups showed the highest performance in adversative connectives and the lowest performance in temporal connectives. Adults performed similarly for all connective types. For the performances of different temporal connectives, all groups (including the adult group) performed better with temporal connective *sonrasında* (after that) than with temporal connective *öncesinde* (before that). Finally, all groups performed better in utterances with subject-referring null pronouns than with object-referring null pronouns, which suggested that null pronouns referring to objects posed a higher challenge. This finding suggested that the pattern of linking null pronouns to subject antecedents more (Şen, 2019; Turan, 1995) might also create a tendency for Turkish readers to interpret null pronouns initially as subject anaphors.

The first striking result of the study was the adult-like performance of HA 5th graders. Unlike the study of Cain and Nash (2011), 5th graders did not significantly differ from adults in overall connective scores. One possible reason is the shallow orthography of Turkish leading to good proficiency in reading at the end of the first formal school year (e.g. Öney and Durgunoğlu, 1997; Öney and Durgunoğlu, 1999). Without the challenges of a deep orthographical language, Turkish children might have improved their overall reading skills along with their discursive awareness (e.g. cohesion and coherence), which then might have helped them to reach adult-like performances in using connectives earlier than reported for other languages. The 3rd graders in the study also showed high performances, even though the overall performance of all 3rd graders was not yet adult-like.

Despite using a standard test for both 5th grader groups, the significant difference between their scores is an interesting finding. The teachers of the middle-school defined 5th grade classrooms as either high-achievers or low-

achievers by considering overall academic performances. Perhaps, the slower acquisition of native language and reading skills, such as connective use, might also interfere with overall academic success (also see Im-Bolter, Yaghoub Zadeh, & Ling 2013). The relationships among academic performance, native language proficiency, and connective acquisition should be further investigated in future studies.

Another possibility for the different findings between this study and Cain and Nash's study might be the test design. Cain and Nash (2011) used three options and a cloze test format. Such designs might have exerted an additional cognitive load especially for younger learners (e.g. 3rd graders). Just as Ercikan and Julian (2010) argued that alternative choice tests with 2 options would be reliable in long tests, the test in this study was highly reliable for 3rd graders (Cronbach's Alpha: 0,865). As LA 5th graders (average age: 10.74) failed to show adult-like performance, it can be suggested that cognitive maturation is not the sole factor in connective acquisition. Not only were LA 5th graders similar to 3rd graders in overall connective test scores, but they also failed to outperform 3rd graders in any of the connective types.

As for the connective types, the children in all age groups performed most successfully with adversative connectives, while adults had no significant differences among their connective type scores. Although some studies have reported extra processing costs for adversatives (Goldman and Murray, 1992; Murray, 1997; Köhne and Demberg, 2013), the highly-constraining nature of adversatives leading to a better understanding (Murray, 1995) explains our findings better. These findings also comply with the findings of Cain and Nash (2005). The nature of the task might also be important in this constraining advantage of adversatives: Both the present study and Cain and Nash's (2005) study used multiple-choice items, in which there were blanks in the middle of sentences. Thus, two contradicting clauses with a blank between them might have indicated an adversative connective more easily.

Temporal connectives were the worst performed connective type for 3rd graders and LA 5th graders. Although one may be tempted to attribute this finding to the maturation factors by referring to the significant differences between the 3rd graders and the adults, it is important to remind that LA 5th graders did not differ from 3rd graders in their temporal connective performance. Thus, individual differences also seem to play a significant role in connective acquisition. To investigate the role of polarity, we compared the scores of the temporal connectives *öncesinde* (before that) and *sonrasında* (after that). All groups, including the adult group, performed worse with the backward temporal connective *öncesinde*. The effect sizes of these differences were especially large for 3rd graders and LA 5th graders (d=1.03 and d=0.89 respectively), indicating that these groups had the greatest difficulty with the questions requiring the connective *öncesinde* as the correct answer.

Considering that the connective *öncesinde* was used as a backward temporal connective to indicate negative polarity in the test, it is likely that high memory demands were responsible for these findings (Crain et al.,1990). Backward temporal connectives might also be contradicting with people's expectations, leading to higher processing costs as the causality-by-default hypothesis suggests (Sanders, 2005).

One of the final findings of the present study was the more difficulty caused by object-referred second clauses compared to subject-referred second clauses. All the second clauses in the study used null pronouns, which referred to objects and subjects in equal numbers. As all the groups in the study did significantly better with the questions including subject-referring null pronouns in their second clauses, the results indicated that the participants were likely to interpret null pronouns as referring to subject initially, which is in line with Turan (1996). Again, 3rd graders and LA 5th graders had more difficulty with object-referred clauses compared to other groups since they showed the largest effect sizes in the comparisons between subject-referred and object-referred clauses (d=0.90 and d=0.75 respectively).

All in all, only HA 5th graders showed an adult-like performance for overall connective scores. As adults showed equal performance for all connective types, and HA 5th graders followed a close pattern, the finding that 3rd graders and LA 5th graders having significant differences among all their connective type scores indicates a developmental pattern. According to the cumulative complexity account of Bloom, temporal connectives should be acquired before adversative and causative connectives; yet, that is contradictory to our findings. 3rd graders and LA 5th graders performed worst with the temporal connectives, while they performed best with the adversative connectives. The two groups also scored significantly lower in the overall connective scores than the other groups while showing the largest effect sizes against the challenges of a backward temporal connective (öncesinde) and object-referred clauses. All those patterns point to an incomplete acquisition of connectives, in which temporal connectives are acquired last. Since the connective *öncesinde* has a negative polarity, it might present an extra challenge in connective acquisition. Evers-Vermeul and Sander's multi-dimensional account explains the current findings better as it considers multiple factors such as conceptual complexity and polarity; rather than depending on semantic relationships, each connective must be considered in terms of its conceptual complexity, polarity, and other characteristics that might be potential factors.

5 Conclusion

This study investigated the acquisition of connectives in Turkish. The results showed that while 3rd graders could comprehend connectives with high-accuracy,

high-achieving (HA) 5th graders were likely to have an adult-like competence already. Individual differences were also an explanatory factor as a group of low-achieving (LA) 5th graders showed a similar performance to 3rd graders. All child groups performed better with adversatives -probably due to the high-constraining nature of this connective type. Correct answer rates were lower for the questions including object-referred second clauses and the questions requiring a backward temporal connective as the correct answer in all groups (even in adults); although these challenges were the most difficult for 3rd graders and LA 5th graders. This study supports a multi-dimensional model of connective acquisition (Evers-Vermeul & Sanders, 2009), which considers multiple factors such as conceptual complexity and polarity.

The present study had some limitations that further studies might avoid. The first was the convenient selection of one of the schools due to practical reasons; to make up for this, classrooms were randomly selected within this school. Also, although dividing the 5th graders into two groups gave us valuable insights regarding the individual differences within the same age group, doing it according to teacher suggestions might be a limitation. Further studies could group children according to their performances in standardized measures of grammar, vocabulary, and working memory to get a more extensive and definitive picture of the individual difference phenomenon in connective acquisition. Finally, although the current task resulted in better performance for adversatives, future studies should further investigate this facilitation effect with different types of tasks.

References

Aksu, A. (1978). The Acquisition of Causal Connectives in Turkish. *Papers and Reports on Child Language Development*, *15*, 129-139.

Aksu-Koç, A. & Slobin, D. (1985). The acquisition of Turkish. *The Crosslinguistic Study of Language Acquisition, Vol.1*. (D.I.Slobin, Ed.). Lawrence Erlbaum, New Jersey

Altenberg, B. (2007). The correspondence of resultive connectors in English. *Nordic Journal of English Studies*, *6*(1).

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language*, *7*(2), 1–21. https://doi.org/10.1017/S0305000900002610

Blything, L. P., Davies, R., & Cain, K. (2015). Young children's comprehension of temporal relations in complex sentences: The influence of memory on performance. *Child Development*, *86*(6), 1922–1934. https://doi.org/10.1111/cdev.12412

Cain, K. E., Patson, N., & Andrews, L. (2005). Age- and ability-related differences in young readers' use of conjunctions. *Journal of Child Language*, *32*(4), 877. https://doi.org/10.1017/S0305000905007014

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*(2), 429–441. https://doi.org/10.1037/a0022824

Crain, S., Shankweiler, D., Macaruso, P. & Bar-Shalom, E. (1990). Working memory and comprehension of spoken sentences: investigations of children with reading disorder. In V. Giuseppe & T. Shallice (Eds), *Neuropsychological impairments of short-term memory* (p. 477-508). New York, NY : Cambridge University Press.

Durgunoğlu, A. Y., & Öney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing*, *11*(4), 281-299.

Ercikan, K., & Julian, M. (2010). Applied Measurement in Education Classification Accuracy of Assigning Student Performance to Proficiency Levels: Guidelines for Assessment Design Classification Accuracy of Assigning Student Performance to Proficiency Levels: Guidelines for Assessment D. Compute, 7347(April 2011), 37–41. https://doi.org/10.1207/S15324818AME1503

Evers-Vermeul, J., & Sanders, T. (2009). The emergence of Dutch connectives; how cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language*, *36*(36), 829–854. https://doi.org/10.1017/

Evers-Vermeul, J., & Sanders, T. (2011). Discovering domains–On the acquisition of causal connectives. *Journal of Pragmatics*, *43*(6), 1645-1662.

Ford, C. E., & Mori, J. (1994). Causal markers in Japanese and English conversations: A cross-linguistic study of interactional grammar. *Pragmatics*, *4*(1), 31–61. https://doi.org/10.1075/prag.4.1.03for

George, D. (2011). *SPSS for windows step by step: A simple study guide and reference, 17.0 update, 10/e*. Pearson Education India.

Goldman, S. R., & Murray, J. D. (1992). Knowledge of connectives as cohesion devices in text: A comparative study of native-English and English-as-a-second-language speakers. *Journal of Educational Psychology*, *84*(4), 504.

Im-Bolter, N., Yaghoub Zadeh, Z., & Ling, D. (2013). Early parenting beliefs and academic achievement: The mediating role of language. *Early Child Development and Care*, *183*(12), 1811-1826.

Köhne, J., & Demberg, V. (2013). The time-course of processing discourse connectives. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, No. 35).

Maury, P., & Teisserenc, A. (2005). The role of connectives in science text comprehension and memory. *Language and Cognitive Processes*, *20*(3), 489–512. https://doi.org/10.1080/01690960444000151.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. transcription format and programs* (Vol. 1). Psychology Press.

Murray, J. D. (1995). Logical connectives and local coherence. *Sources of coherence in reading*, 107-125.

Murray, J. D. (1997). Connectives and narrative text: The role of continuity. *Memory & Cognition*, *25*(2), 227-236.

Öney, B., & Durgunoğlu, A. Y. (1997). Beginning to read in Turkish: A phonologically transparent orthography. *Applied Psycholinguistics*, *18*(1), 1-15.

Sanders, T., Land, J., & Mulder, G. (2007). Linguistics markers of coherence improve text comprehension in functional contexts. *Information Design Journal*, *15*(3), 219–235. https://doi.org/10.1075/idj.15.3.04san

Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical methods in medical research*, *8*(1), 3-15.

Sofu, H., & Şimşek, T. (2016, February) Acquisition of Coordination in Turkish Children: Additive Connectives. *Paper presented at the 3rd International Conference on Language, Innovation, Culture and Education (ICLICE),* Singapore.

Şen, B. (2019). *An Analysis of Turkish Singular Pronouns Based on Accessibility Theory* (Master's thesis, Hacettepe University, Ankara, Turkey). Retrieved from http://www.openaccess.hacettepe.edu.tr:8080/xmlui/bitstream/handle/11655/7714/10262139.pdf?sequence=1&isAllowed=y

Torabi Asr, F., & Demberg, V. (2012). Implicitness of discourse relations. *Coling*, (December 2012), 2669–2684.

Turan, Ü. D. (1996). Null vs. overt subjects in Turkish discourse: A centering analysis. *IRCS Technical Reports Series*, 95.

van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*(4), 584-589.

van Silfhout, G., Evers-Vermeul, J., & Sanders, T. (2014). Connectives as processing signals: How students benefit in processing narrative and expository texts. *Discourse Processes*, *52*(1), 47–76. https://doi.org/10.1080/0163853X.2014.905237

Zufferey, S., Mak, W. M., & Sanders, T. J. M. (2015). A cross-linguistic perspective on the acquisition of causal connectives and relations. *International Review of Pragmatics*, *7*(January), 22–39. https://doi.org/10.1163/18773109-00701002