

# Acquiring Persian Object Marking: Balochi Learners of L2 Persian

Ute Bohnacker, Uppsala University

ute.bohnacker@lingfil.uu.se

Somayeh Mohammadi, Azad University of Bam

somayehmohammadi77@yahoo.com

## Abstract

This paper investigates the second language acquisition of Persian object marking by 237 native Balochi-speaking children from age eight to eleven. The combination of Balochi and Persian has not been studied from an acquisitional perspective, although it widely occurs in southeastern Iran. This paper studies the second language (L2) development of direct and indirect object marking after two and three years of exposure to Persian at school, vs. a control group of 133 monolingual native Persian children of similar age and socio-economic background. Beginning learners predominantly use non-targetlike object marking constructions reminiscent of Balochi, whilst there is a clear trend towards Persian-style object marking in the more advanced learners. These findings suggest that L2 learners initially transfer core aspects of their native grammar to their interlanguage Persian. The picture is complicated by the existence of differential object marking in Balochi and Persian and the somewhat different writing patterns of native and non-native learners.

*Keywords:* Balochi, Persian, children, L2 acquisition, object marking, *rā*, differential object marking (DOM), specificity, written composition, L1 transfer, preposition, indirect object, direct object, discourse pragmatics

## 1. Introduction

This paper investigates the second language (L2) acquisition of Persian direct and indirect object marking by children with Balochi as their first language (L1). Balochi is a northwestern Iranian language closely related to Persian, but differing from it in several aspects of morphology and syntax. For example, Persian marks indirect objects with a preposition and direct objects postnominally, whilst Balochi employs one and the same suffix on indirect and direct objects. Both languages exhibit differential object marking, which means that the overtness of object markers is influenced by semantic and discourse-pragmatic factors. Moreover, the overtness of object markers in Balochi, but not in Persian, is influenced by split ergativity.

The language pair Balochi/Persian has not been studied empirically from an acquisitional perspective, although it widely occurs in the southeastern provinces of Iran where Balochi is the native language but schooling takes place in Persian only. A pilot study found Persian object marking to be a particularly problematic area for learners. The present paper investigates this issue more systematically, by studying the development of indirect and direct object marking in the L2 Persian writing of 237 Balochi L1 school children aged eight to eleven. Two groups of learners with different lengths of L2 exposure are compared with each other and also with a control group of 133 monolingual L1 Persian children of similar age and socio-economic background. Large differences between learner groups are found. For the beginning L2 learners, non-targetlike indirect and direct object marking constructions

predominate, most of which are reminiscent of Balochi, whereas the more proficient learners show a clear trend towards targetlike, Persian-style object marking. These findings suggest that the elementary L2 learners transfer core aspects of their L1 Balochi grammar to their interlanguage Persian, whilst only vestiges of such transfer remain in the productions of the more advanced learners. This picture is somewhat complicated by the existence of differential object marking in Balochi and Persian and by slightly different writing patterns in native and non-native learners.

The present paper is structured as follows. In order to familiarize readers with the morpho-syntax of object marking in the two languages, Section 2 describes the Persian constructions of indirect object marking and direct object marking, including differential object marking, in light of the existing literature. Section 3 then outlines the Balochi object marking constructions, based on scholarly publications and our own observations. Section 4 summarizes the little research there is on the acquisition of Persian object marking. Section 5 provides background and methodological information on our current study. Section 6 analyses, compares and discusses the object marking results in the beginning L2 learners, more advanced L2 learners, and native Persian controls, concerning overall frequencies (6.1), indirect objects (6.2), and direct objects (6.3). Section 7 concludes the paper.<sup>1</sup>

## 2. Persian

Persian is a null-subject language with rich verbal morphology and subject-object-verb (SOV) word order. The only case marker is postnominal *rā* (discussed below) which signals the grammatical function of direct object. Subjects are not marked for case, and roles of arguments other than subject or direct object are marked by prepositions (see below). In nominal expressions, modifiers typically follow the nouns they modify. Nouns and pronouns are invariant in form in Persian. There is no grammatical gender and nominal morphology is limited to number marking (singular vs. plural) and optional indefinite marking.<sup>2</sup>

<sup>1</sup> We would like to thank Carina Jahani for many useful comments and suggestions and for encouraging us to write this paper in the first place. Thanks also to an anonymous *Orientalia Suecana* reviewer for suggestions.

<sup>2</sup> In Persian, (in)definiteness is not always morphologically marked, and the interpretation of noun phrases as indefinite or definite is often context-dependent. For instance, *sib xarid-i?* (apple bought.PAST-2SG) with the bare noun *sib* can mean ‘Have you bought apples?’ in an all-new context, but also ‘Have you bought the apples?’ in a context where apples or apple-buying had been discussed before (cf. Rahimian and Hajiani 2009: 400). Indefiniteness, if overtly marked, can be indicated by prenominal *yek* ‘one, a’ in the singular, or by indefinite enclitic *i* (‘a, some’), either on its own or in combination with *yek*. (Thus *ketāb-i*, *yek ketāb* and *yek ketāb-i* all mean ‘a book’.) Unlike *yek*, *i* may combine with both singular and plural nouns (e.g. Lazard 1992; Lyons 1999; Ghomeshi 1996, 2003; Ganjavi 2007: 142–143 fn. 67). The enclitic *i* is mobile and can attach either to the noun or to an adjective if there is one in the noun phrase: *ketāb-i gerān* / *ketāb-e gerān-i* ‘an expensive book’ (*-e* is a linking vowel here). Definiteness tends not to be morphologically marked in Persian. There is no definite article, and bare noun phrases allow specific, definite interpretations (e.g. *lebās* ‘the dress’). A demonstrative (*ān/un/in* ‘that/this’) can be added for deixis and definiteness, and in colloquial speech, the optional enclitic *e* may appear to signal definiteness (e.g. *man un lebās-e-rā xarid-am* ‘I that dress-DEF-OBJ bought.PAST-1SG ‘I bought the dress.’, cf. Samian 1983; Ganjavi 2007:58; Keshavarz 2007: 259). Such definite *e* does not, however, appear in writing or in the more formal varieties of spoken Persian.

### 2.1. Indirect object marking in Persian

In Persian, one preposition, *be* ‘to’, is used to signal the grammatical function of indirect object, which often expresses the role of goal or beneficiary (e.g. to tell someone, to give someone something). *Be* is used with all indirect objects, whether they are noun phrases headed by a proper noun, common noun, or pronoun, as illustrated in (1)–(4). This holds for spoken and written present-day Persian alike. *Be* is used irrespective of whether the indirect object has specific reference (as in (1)–(3)), un-specific or generic reference (as in (4)).

- (1) Per. (man) {*be Sārā/be moʔalem/be mādar-am/be u*} *goft-am*  
 I to Sara/to teacher/to mother-my/to (s)he told.PAST-1SG  
 ‘I told Sara/the teacher/my mother/her.’
- (2) Per. *u* {*be Sārā/be dokhtar/be u/be man*} *pul mi-dah-ad*  
 (s)he to Sara/to girl/to (s)he/to I money IMPF-give.PRES-3SG  
 ‘She gives Sara/the girl/her/me money.’
- (3) Per. *quri-rā* {*be Sārā/be mādar-am/be u*} *dād-am*  
 teapot-OBJ to Sara/to mother-my/to (s)he gave.PAST-1SG  
 ‘I gave Sara/my mother/her the teapot.’
- (4) Per. *doktor-hā be mardom komak mi-kon-and*  
 doctor-PL to people help IMPF-do.PRES-3PL  
 ‘Doctors help people.’

### 2.2. Direct object marking in Persian

Persian direct objects are not marked by a preposition but by postnominal *rā*, sometimes referred to as a suffix, sometimes referred to as a postposition or enclitic particle (e.g. Bossong 1985; Windfuhr 1979: 47–57; Dabir-Moghaddam 1990; Lazard 1992 [1957]: 74–76, 183–193; Karimi 1990, 1996). *Rā* in its written form is رَا, which is pronounced [rɑ] in more formal varieties, and [ro], [rə] or [o] or [ə] in more colloquial style, depending on whether it attaches to a word that ends in a vowel or a consonant (Lazard 1970: 74; Karimi 1990: 139; Ganjavi 2007: 9, 108). In writing, *rā* appears separately or joined to the preceding word. *Rā* is used in both written and spoken varieties of Persian, and it appears on direct objects irrespective of whether they are headed by a proper noun, common noun, or pronoun, as shown in (5)–(6). When the direct object consists of a more complex noun phrase, such as one containing an adjective or a possessive enclitic, *rā* attaches to the end of the phrase, as in (7)–(8).

- (5) Per. (man) {*Sārā-rā/moʔalem-rā/mādar-am-rā/u-rā*} *did-am*  
 I Sara-OBJ/teacher-OBJ/mother-my-OBJ/she-OBJ saw.PAST-1SG  
 ‘I saw Sara/the teacher/my mother/her.’
- (6) Per. *ki-rā* *did-i*  
 who-OBJ saw.PAST-2SG  
*Sārā-rā/moʔalem-rā/mādar-am-rā/u-rā*  
 Sara-OBJ/teacher-OBJ/mother-my-OBJ/she-OBJ  
 ‘Who did you see? – I saw Sara/the teacher/my mother/her.’

- (7) Per.      *quri-ye bozorg-rā*                      *kharid-am*  
 teapot-LINK big-OBJ                      bought.PAST-1SG  
 ‘I bought the big teapot.’
- (8) Per.      *mādar-e Leilā-rā*                      *did-am*  
 mother-LINK Leila-OBJ                      saw.PAST-1SG  
 ‘I saw Leila’s mother.’

However, not every direct object in Persian is marked by *rā*, neither in the written nor in the spoken modality. As illustrated in (9)–(10), *rā* is obligatory on proper nouns and personal and demonstrative pronouns, which are inherently definite and specific.

- (9) Per.      *Sārā\*(-rā)*                      *did-am*  
 Sara-OBJ                      saw.PAST-1SG  
 ‘I saw Sara.’
- (10) Per.      *u\*(-rā)*                      *did-am*  
 (s)he-OBJ                      saw.PAST-1SG  
 ‘I saw him/her.’

For direct objects other than proper nouns and pronouns, i.e. noun phrases headed by a common noun, there is variability in the use of *rā*. This variability is also known as differential object marking, where a number of semantic and discourse-pragmatic factors influence the realization or non-realization of object marking (e.g. Lazard 1992: 183–194; Bossong 1985: 3, 57–67; Windfuhr 1989: 533, 1992: 31).

Linguists are not in agreement as to which factors contribute most to *rā* being realized on the direct object in Persian. Both intrinsic semantic properties of the argument (humanness, animacy) and extrinsic semantic and discourse-pragmatic factors, such as information status (e.g. identifiability) and referentiality (definiteness, specificity) may be involved (cf. Browne 1970; Lazard 1982, 1984, 1992; Windfuhr 1989: 533; Karimi 1990, 1996, 2003; Dabir-Moghaddam 1990: 32–35; Ghomeshi 1996, 1997, 2003; Shokouhi and Kipka 2003; Ganjavi 2007: 109–113, 142–150; Rahimian and Hajiani 2009). Accounts of differential object marking in Persian differ from publication to publication, and the authors’ introspective grammaticality judgments often diverge for decontextualized examples with or without *rā* (see Shokouhi and Kipka 2003: 953–957 for an overview).

Lazard (1982, 1984, 1992) voiced the view that extrinsic semantic factors (definiteness, specificity) play a larger role than humanness or animacy in Persian. According to him, definite objects are virtually always *rā*-marked, regardless of animacy; specific indefinites are, regardless of animacy, occasionally *rā*-marked; and nonspecific indefinites/generics are normally not *rā*-marked. For the latter, Lazard suggests that the use of *rā* is influenced by the animacy of the object; *rā* is preferred for humans and dispreferred for inanimates (Lazard 1982: 181–186; 1984: 278–283; 1992: 185). These claims have not yet been tested empirically via corpus studies or experimental data. Lazard’s idea that definiteness and specificity are the main fac-

tors behind *rā* marking has however led scholars to claim that *rā* exclusively appears on definite objects and functions as a definiteness marker, or that *rā* exclusively appears on specific objects and functions as a specificity marker (cf. e.g. Browne 1970: 362; Comrie 1989: 132–135; Karimi 1989, 1990). Specificity here refers to the selection of a particular individual from a set of individuals (Karimi 1990: 142–145). Other scholars have however pointed out examples where direct objects with a specific reading occur without *rā*, or alternatively where *rā*-marked objects allow indefinite and/or nonspecific interpretations (e.g. Dabir-Moghaddam 1990; Shokouhi and Kipka 2003; Rahimian and Hajjani 2009).<sup>3</sup>

We will not go further into this debate here, but simply treat *rā* as an object marker. Based on our understanding of the literature and our own observations of Persian, we believe it is fair to describe the general tendency concerning the use of Persian *rā* on direct objects as follows: The higher the argument is on the animacy scale (e.g. + human) and/or the more identifiable the argument is in context (and note that identifiability often goes together with specificity and definiteness) the more likely it is that it will carry the object marker *rā*.<sup>4</sup>

For instance, on direct objects denoting specific humans like Sara or the teacher or she as in (8)–(10) above and (11) below, *rā* is basically obligatory.

- (11) Per.    *moʔalem\*(-rā)*            *did-am*  
               teacher-OBJ                saw.PAST-1SG  
               ‘I saw the teacher.’

By contrast, the less specific, the less definite, the less given and identifiable a direct object is in context, the less likely it will be that *rā* is used. For instance, when expressing the notion of having children, as in (12), the direct object *bache* ‘children’, whilst denoting humans, does not refer to any identifiable, specific children in context. In such a case, *rā* is not used. Similarly, when expressing the notion of giving money to the poor, as in (13), the direct object *pul* ‘money’ does not refer to any identifiable, specific money in context. Again, *rā* is not used.

- (12) Per.    *man*        *bache(\*-rā)*        *dar-am*  
               I            child                have.PRES-1SG  
               ‘I have children.’

- (13) Per.    *man*        *be fogharā*        *pul(\*-rā)*        *mi-dah-am*  
               I            to poor.PL        money            IMPF-give.PRES-1SG  
               ‘I give money to the poor.’

<sup>3</sup> Moreover, *rā* cannot be a generalized definiteness or specificity marker, since arguments that are not direct objects, i.e. subjects, predicative nominals, and objects of prepositions, do not carry *rā* even if they are specific (Karimi 1990).

<sup>4</sup> Some authors working within the generative Minimalist framework try to derive the distributional realization of *rā* from the abstract internal syntactic structure of Persian nominal expressions. One such approach is Ganjavi (2007), who assumes that only DPs (Determiner Phrases), i.e. nominal expressions with what she regards as a complete setup of functional projections, are *rā*-marked, whilst other nominal expressions are seen as lacking some abstract functional projections and therefore lack *rā* marking as well. In the present paper, we will not discuss whether *rā* marking on direct objects should be seen as a by-product of the abstract syntactic structure of nominal expressions, as this is influenced by the author’s predilection and choice of syntactic model.

The use of *rā* on the direct object is influenced by the identifiability of the object in context. Thus, when expressing the notion of buying a house and referring to a specific, particular and identifiable house as in (14a), *rā* must be used on *khāne* ('the house'); whilst *khāne* without *rā* is the preferred way of expressing 'a house, an un-specific house' as in (14c).<sup>5,6</sup>

- (14) a. *mi-khāh-am*                      *khāne\*(-rā)*                      *be-khar-am*  
 IMPF-want.PRES-1SG                      house-OBJ                      SBJ-buy.PRES-1SG  
 'I want to buy the house.' (a specific house which the speaker assumes to be known/identifiable by the listener)
- (14) b. *mi-khāh-am*                      *khāne-i-rā*                      *be-khar-am*  
 IMPF-want-1SG                      house-INDEF-OBJ                      SBJ-buy.PRES-1SG  
 'I want to buy a (certain) house.' (a specific house known to the speaker but assumed to be unknown/not identifiable by the listener)
- (14) c. *mi-khāh-am*                      *khāne (\*-rā)*                      *be-khar-am*  
 IMPF-want-1SG                      house                      SBJ-buy-1SG  
 'I want to buy a house.' (some un-specific house as yet unknown)

Publications on *rā* are typically based on unquantified decontextualized single-sentence examples. We have not been able to find any studies that could inform us about the frequency and extent to which direct objects are marked with *rā* in Persian. This would be most illuminating, both concerning different modalities (spoken, written) and different genres. One rare corpus study is Shokouhi and Kipka (2003), who analysed six hours of colloquial spoken dialogue between young expatriate Persian-speaking academics in Australia (5000 intonation units). They found 233 instances of direct object marking with *rā*, and the authors' breakdown by information type shows that the majority of these tokens were informationally given or accessible referents (88%). 12% of the tokens with *rā* encoded new referents, but even these overwhelmingly turned out to be identifiable in context (2003: 958–962). Shokouhi and Kipka also looked at all direct objects that were informationally new in the context they occurred in and found that 87% of these objects were not marked with *rā*. This suggests that there is a correlation between identifiability and overt *rā* marking of the direct object, but not a one-to-one correspondence (2003: 962). However, as Shokouhi and Kipka (2003) did not investigate all the direct objects without *rā*, their study cannot provide any information on the frequency of *rā* marking in Persian as such.

<sup>5</sup> *Rā* is not restricted to definites, but can cooccur on direct objects with the indefinite postnominal *i*, as in (14b) and/or with the indefinite prenominal *yek* 'one, a'. According to Browne (1970: 361) and Ghomeshi (2003), such overtly marked indefinite direct objects with *i* and *rā* are interpreted as identifiable, specific indefinites.

<sup>6</sup> There are two types of *rā* constructions in Persian. Apart from direct object marking, *rā* can also be used to mark sentence-initial aboutness topics of the type "As for X, ...". As a topic marker, *rā* can occur on elements that are not direct objects, and as a consequence, there may be sentences that contain two instances of *rā*, one a topic marker and one a direct object marker. Persian topic marking with *rā* will not be dealt with any further here (for discussion see e.g. Karimi 1990: 143–158; Shokouhi and Kipka 2003; Ganjavi 2007: 106–109, 135–166).

### 3. Balochi

Balochi is a northwestern Iranian language closely related to Persian. It is the principal language of the Baloch of Balochistan in Iran, Pakistan, and southern Afghanistan. Balochi does not have a standardized language or a standardized writing system, and it mostly remains an oral language, consisting of several regional dialects. Education in Balochi-speaking areas invariably takes place in a second language such as Persian or Urdu, which means that Balochi is largely restricted to the informal language domains of the home and neighbourhood and to traditional occupations (Jahani and Korn 2009: 635). We will here concern ourselves only with the Balochi variety spoken in Iranshahr (in the Sistan and Balochistan province of southeastern Iran), because this is where our empirical study on Balochi learners of Persian was carried out. Descriptions of Balochi can be found in Jahani (2003) and Jahani and Korn (2009) and references cited therein.

Like Persian, Balochi is a null-subject language with rich verbal morphology and subject-object-verb word order. Many dialects of Balochi have split ergativity. In the present/future tense (i.e. non-past), a nominative-accusative system is used, whilst the past tense has an ergative system. Consequences of this split for object marking are discussed below.

In Balochi nominal expressions, attributive adjectives, possessives, and quantifiers precede the nouns they modify. Balochi has no grammatical gender, and nominal morphology is limited to number marking (singular/plural), optional indefinite marking,<sup>7</sup> and some case-marking. In the nominative-accusative system, the suffix *-a* (discussed below) often occurs to signal the grammatical function of object, whilst subjects are not overtly marked (Jahani 2003: 118).

#### 3.1. Indirect and direct object marking in Balochi as compared to Persian

Balochi uses the same object marker on both indirect and direct objects, and not only on direct objects as in Persian. Object marking in Balochi may thus be likened to some older versions of Persian, where *rā* marking did occur on both indirect and direct objects (e.g. Dabir-Moghaddam 1990: 32 for Early Modern Persian). In present-day Persian, object marking with *rā* is however restricted to direct objects. There are some prepositions in Balochi but, in contrast to Persian, they are not used to mark indirect objects. The Balochi object marker *-a* has a number of allomorphs, *-a*, *-ya*, *-ra*, depending on whether it is preceded by a consonant or a vowel (Windfuhr 1992: 31; Jahani 2003: 115, 118; Jahani and Korn 2009: 651–652). Indirect objects are illustrated in (15)–(16), direct objects in (17)–(18). As shown in the examples, the morphological marking on indirect and direct objects is identical.

<sup>7</sup> Indefiniteness can be marked on Balochi noun phrases by enclitic *e* ('a'), which attaches to nouns in the singular (e.g. *kitab-e* 'a (certain) book'), or by the pronominal indefinite article/numeral, *ya* 'a, one'. *Ya* and *e* can also co-occur (e.g. *ya kitab-e* 'a book'). According to Jahani and Korn (2009: 667), overt marking with *e* induces an indefinite specificity reading ('a certain book').

## Balochi indirect objects

- (15) Bal. {*Sārā-ya/mallem-a/wtī māt-a/āyī-a*} *gwash-on*  
Sara-OBJ/teacher-OBJ/my mother-OBJ/she-OBJ tell.PRES-1SG  
'I tell Sara/the teacher/my mother/her.'
- (16) Bal. *āyī* {*Sārā-ya/mallem-a/wtī māt-a/āyī-a/man-a*} *zarr-a dant*  
(s)he Sara-OBJ/teacher-OBJ/my mother-OBJ/she-OBJ/ I-OBJ money-IMPF  
give.PRES.3SG  
'She gives Sara/the teacher/my mother/her/me money.'

## Balochi direct objects

- (17) Bal. *kay-a* *gend-e*  
who-OBJ see.PRES-2SG  
{*Sārā-ya/mallem-a/wtī māt-a/āyī-a*} *gend-on*  
Sara-OBJ/teacher-OBJ/my mother-OBJ/ see.PRES-1SG  
(s)he-OBJ  
'Who do you see? – I see Sara/the teacher/my mother/her.'
- (18) Bal. *Maryam* {*jāmag-a/ketāb-a*} *gī*  
Maryam dress-OBJ/book-OBJ buy.PRES.3SG  
'Maryam buys the dress/the book.'

Readers may have noticed that the above examples of object marking in Balochi (15)–(18) are exclusively in the present tense. In all tenses formed with the present stem, sentences in Balochi are constructed following a nominative-accusative system, and here overt object marking with *-a* is used. In the past tense, however, sentences are constructed ergatively and direct objects are not marked overtly (Korn 2009; Jahani and Korn 2009: 669). This is illustrated by the following minimal-pair examples contrasting Balochi and Persian, (19) for non-past, and (20) for past tense.

(19) Present tense: overt object marking in both Balochi and Persian<sup>8</sup>

- a. Bal. {*Sārā-ya/mallem-a/wtī māt-a/āyī-a*} *gend-on*  
Sara-OBJ/teacher-OBJ/my mother-OBJ/(s)he-OBJ see.PRES-1SG
- a'. Per. {*Sārā-rā/moʔalem-rā/mādar-am-rā/u-rā*} *mi-bin-am*  
Sara-OBJ/teacher-OBJ/mother-my-OBJ/(s)he-OBJ IMPF-see.PRES-1SG  
'I see Sara/the teacher/my mother/her.'
- b. Bal. *āyī* {*Sārā-ya/mallem-a/wtī māt-a/āyī-a*} *zarr-a dant*  
(s)he Sara-OBJ/teacher-OBJ/my mother-OBJ/(s)he-OBJ money-IMPF  
give.PRES.3SG
- b'. Per. *u* {*be Sārā/be moʔalem/be mādar-am/be u*} *pul mi-dah-ad*  
(s)he to Sara/to teacher/to mother-my/to (s)he money IMPF-  
give.PRES-3SG  
'(S)he gives Sara/the teacher/my mother/her money.'

<sup>8</sup> Balochi person and number markings on the verb (as *-on* in (19a)) are often homophonous with pronominal clitics (CL) that attach to other elements such as noun phrases (as *-(y)on* in (20a)). Pronominal clitics often specify the person and number of the agent (Jahani and Korn 2009: 654; Korn 2009: 57–60).

(20) Past tense: no overt direct object marking in Balochi, but overt marking in Persian

- |          |   |               |
|----------|---|---------------|
| a. Bal.  | { <i>Sārā-yon/mallem-on/wtī māt-on/āyī-on</i> }   | <i>di(st)</i> |
|          | Sara-1SG.CL/teacher-1SG.CL/my mother-1SG.CL/<br>(s)he-1SG.CL                            | saw.PAST.3SG  |
| a'. Per. | { <i>Sārā-rā/mo?alem-rā/mādar-am-rā/u-rā</i> }  | <i>did-am</i> |
|          | Sara-OBJ/teacher-OBJ/mother-my-OBJ/(s)he-OBJ<br>'I saw Sara/the teacher/my mother/her.' | saw.PAST-1SG  |

There is thus an important difference concerning object marking between Balochi and Persian: Direct object marking can be overt in all tenses for Persian, but in Balochi, due to split ergativity, object marking is only overt for tenses formed from the present stem. Irrespective of tense, there are also some other ergative constructions in Balochi where a pronominal clitic (e.g. *-on*, *-ī*) attaches to the direct object whilst the object marker *-a* vanishes, as in (20) above.<sup>9</sup> We need to keep this in mind when considering the acquisition of Persian object marking by Balochi pupils.

In connection with direct object *rā* in Persian, we previously discussed differential object marking (DOM). Like Persian, Balochi also exhibits DOM, and thus *-a* on the object can be realized or dropped depending on certain semantic and discourse-pragmatic factors. DOM has not been studied systematically for Balochi (Farrell 1990: 65; Korn 2009), and there are no corpus (frequency) studies of the language as yet. However, Jahani and Korn (2009: 669–670) cite examples from a number of Balochi dialects that suggest that overt *-a* marking on the direct object goes together with specificity and context-identifiability of the argument. Generic, non-specific and inanimate objects, on the other hand, tend to be unmarked (Jahani and Korn 2009: 669). Our own informal observations of Balochi (S. Mohammadi) support this; see the examples in (21)–(25). Thus, DOM of direct objects in Balochi appears to pattern similarly to what has been observed for DOM in Persian.

DOM: Overt object *-a* marking on human/animate and specific/context-identifiable objects

- |           |  |                            |
|-----------|--|----------------------------|
| (21) Bal. | <i>man</i> { <i>wtī dōst*(-a)/āyī*(-a)</i> } | <i>har rōč-a gend-on</i>   |
|           | I my friend-OBJ/her-OBJ                      | every day-IMP see.PRES-1SG |
|           | 'I see my friend/her every day.'             |                            |

<sup>9</sup> Consider the following Balochi example (i) from Iranshahr (S. Mohammadi). The direct object carries an object marker *-a* when the agent (*-on*) is marked on the verb (a-version), but this object marker *-a* disappears when the pronominal clitic agent attaches to the object (b-version).

- |             |  |
|-------------|--|
| (i) a. Bal. | <i>Maryam-a dusta dar-on</i>   |
|             | Maryam-OBJ love have.PRES-1SG  |
| b. Bal.     | { <i>Maryam-(*a)-on/wtī mat-(*a)-on</i> } <i>dust-en</i>                   |
|             | Maryam-1SG.CL/my mother-1SG.CL love.PRES-COP<br>'I love Maryam/my mother.' |
| (ii) Per.   | { <i>Maryam-rā/mādar-am-rā</i> } <i>dust dār-am</i>                        |
|             | Maryam-OBJ /mother-my-OBJ love have.PRES-1SG<br>'I love Maryam/my mother.' |

- (22) Bal. *nā-bāyad-ē* *chokk-ān\*(-a)* *aziyat kan-en*  
 NEG-must- child-PL.OBL-OBJ annoy do.PRES-1PL  
 COP.PRES.3SG  
 ‘We mustn’t annoy the children.’

DOM: No overt object *-a* marking on animate, unspecific/indefinite objects

- (23) Bal. *man lōt-on* *ya {aps-ē/bat-ē}* *be-ger-on*  
 I want.PRES-1SG one horse-INDEF/ SBJ-buy.PRES-1SG  
 duck-INDEF  
 ‘I want to buy a horse/duck.’ (some unspecific horse/duck)

DOM: No overt object *-a* marking on non-human/inanimate, unspecific/indefinite objects

- (24) Bal. *man ya davār-ē* *gend-on*  
 I one house-INDEF see.PRES-1SG  
 ‘I (can) see a house.’
- (25) Bal. *man lōt-on* *ya {gūrī-ē/davār-ē}* *be-ger-on*  
 I want.PRES-1SG one teapot-INDEF/ SBJ-buy.PRES-1SG  
 house-INDEF  
 ‘I want to buy a teapot/a house.’ (some unspecific teapot/house)

Concerning indirect object marking with *-a* in Balochi, the literature is sparse, but Jahani and Korn (2009: 670) state that “indirect objects are marked” and our own observations of Balochi (S. Mohammadi) confirm this. In natural discourse, indirect objects are typically human (e.g. to help someone, give something to someone, tell someone), a tendency that is pronounced in colloquial speech and child language. It is thus not surprising that Balochi indirect objects, due to the strong bias of the argument to denote humans, carry overt *-a* marking (in the nominative-accusative system), irrespective of whether they are identifiable in context (e.g. the teacher, (s)he) or non-identifiable (e.g. the poor, people), see the examples in (26)–(29).

- (26) Bal. {*malle\*(-a)/āyī\*(-a)*} *gwasht-on*  
 teacher-OBJ/(s)he-OBJ tell.PAST-1SG.CL  
 ‘I told the teacher/her.’
- (27) Bal. *Sārā āyī\*(-a)* *komak-a kan*  
 Sara (s)he-OBJ help-IMPF do.PRES.3SG  
 ‘Sara helps her.’
- (28) Bal. *man garīb-ān\*(-a)* *pul-a day-on*  
 I poor-PL.OBL-OBJ money-IMPF give.PRES-1SG  
 ‘I give money to the poor.’
- (29) Bal. *doktor-ān mardom\*(-a)* *komak-a kan-an*  
 doctor-PL.NOM people-OBJ help-IMPF do.PRES-3PL  
 ‘Doctors help people.’

In sum, when comparing Balochi and Persian, at first glance the two languages do not appear very different from each other as far as direct object marking is concerned. Both use morpho-phonologically similar markers, *rā* (Persian) and *-a* (Balochi), and both do so in the same position, postnominally. Both languages allow differential object marking (DOM), and although the last word has not been said concerning the factors that determine the realization or non-realization of overt *rā*/*-a*, both intrinsic semantic properties of the noun (humanness, animacy) and discourse-pragmatic features (definiteness, specificity, givenness/identifiability) appear to play a role. Whilst we acknowledge the need for a more systematic study of these factors in the two languages, our impression is that Persian and Balochi do not greatly differ concerning DOM. In both languages, direct objects denoting a human, specific referent typically carry overt marking, irrespective of the type of the nominal expression (proper noun, common noun phrase, personal pronoun). On the other hand, direct objects which denote inanimate, nonspecific, and new-information referents are typically not marked overtly with *rā* (Persian) or *-a* (Balochi). Recall however that due to split ergativity, Balochi overtly marks objects only in the nominative-accusative paradigm.

Concerning the marking of indirect objects in Balochi and Persian, there is a clear morpho-syntactic difference between the two languages: Persian uses the preposition *be* and Balochi the suffix *-a*, though, as will be recalled, Balochi only uses this overt object marking in the nominative-accusative paradigm, particularly in the present tense system.

#### 4. Previous studies of acquisition of object marking in Persian

Monolingual children appear to acquire Persian *rā* marking early on. A recent doctoral thesis (Foroodi-Nejad 2011) on four-to-seven-year-old Persian-speaking children in Iran includes an oral elicitation experiment concerning direct object *rā* marking, amongst other topics. Foroodi-Nejad found that *rā* is mastered relatively early by typically developing children, whilst omission of *rā* from obligatory contexts may be a clinical marker of language impairment in monolingual children. Typically developing monolinguals regularly use *rā* in their speech, and in the experiment they use it in a targetlike fashion in 97% of obligatory contexts already by 4 years (and probably earlier, but 4 years was the youngest age studied, cf. Foroodi-Nejad 2011: 83–103).<sup>10</sup> Unfortunately, there is no corresponding data concerning the acquisition of object marking in bilingual or child L2 learners of Persian.

To our knowledge, object marking in Balochi-speaking second language learners of Persian has not yet been researched. In general, there are few studies of L2 Persian, and there is only one study that we are aware of that deals with the L2 acquisition of Persian object marking (Mirza 2000). In her MA thesis, Mirza looked at the case system in Armenian learners of Persian who had been exposed to Persian since early childhood. In spite of the fact that Armenian and Persian objects are marked

<sup>10</sup> Foroodi-Nejad (2011) only elicited *rā* in obligatory contexts. Therefore, her children's 97% provision of *rā* cannot be taken to mean four-year-olds *rā*-mark 97% of all their direct objects (cf. 2.2.).

very similarly (and both languages have DOM), Mirza found that object marking was a consistent problem for a group of Armenian learners of L2 Persian. In a grammaticality judgment task, untutored L2 learners (illiterate adults) accepted Persian sentences with a direct object with a specific reading without *rā* marking. Native Persian speakers and Armenian adults who had had L2 schooling in Persian and more exposure to Persian rejected the same sentences as ungrammatical, requiring the direct object to be marked with *rā* (as in (5)–(11)).<sup>11</sup> Mirza's untutored L2 group also "corrected" grammatical Persian stimulus sentences with overt *rā* by removing *rā* from direct objects with a specific reading, which resulted in ungrammaticality. By contrast, the tutored Armenian learners of Persian kept overt *rā*. Mirza's study concerns adults only, and we have not been able to find any studies of child L2 Persian with which we could compare the results of our own study. However, Mirza's and Foroodi-Nejad's (2011) studies do suggest that Persian object marking might be a problematic area for certain language learners.<sup>12</sup>

## 5. The current L2 acquisition study: Background and method

This study investigates the acquisition of direct and indirect object marking in Persian as a second language by Balochi-speaking school children in southeastern Iran. We became interested in this issue as all literacy training and school teaching for Balochi children takes place in Persian only, and the children are expected to be able to write in Persian from second grade. In a pilot study, Mohammadi (2009) noted that one area that Balochi school children struggle with is Persian object marking. The present study investigates this issue more systematically by looking at the development of indirect and direct object marking in the L2 Persian compositions of 237 Balochi L1 school children between the ages of eight and eleven. Two groups of learners with different lengths of L2 exposure are compared with each other and also with a group of 133 age-matched monolingual L1 Persian children, all of them writing a composition on the same topic under similar conditions.

The Balochi L1 children that participated in the study had started school at the age of seven and were attending unisex primary schools in an urban area of southeastern Iran (Fatemeh Zahra Primary and Zakiye Primary, both in Mohammadan town, Iranshahr city). The children were all girls. Care was taken to keep their language background as homogeneous as possible. Therefore, only children who came from monolingual Balochi homes were included. Concerning the social background of the children, most came from low literacy homes. The majority of their parents

<sup>11</sup> Here are two of Mirza's examples:

- |      |   |  |
|------|---|--|
| (i)  | <i>Nāhid Maryam(-rā) be kelas āvard</i><br>Nahid Maryam-RA to class brought<br>'Nahid brought Maryam to class.' | (obligatory <i>rā</i> in native Persian, but optional for L2ers) |
| (ii) | <i>Mas'ud be Ali in pul(-rā) dād</i><br>Masood to Ali this money-RA gave<br>'Masood gave this money to Ali.'    | (obligatory <i>rā</i> in native Persian, but optional for L2ers) |

<sup>12</sup> We are also aware that the acquisition of differential object marking in other languages, such as Spanish, is a problematic area for bilingual/L2 learners (see e.g. Bowles and Montrul 2009).

were illiterate; some had attended primary school (age 7–11) and/or middle school (up to age 13), but only a few had gone on to secondary education or received a diploma (the Iranian equivalent of having attended school for 12 years).

The children in the study had only sporadically come in contact with languages other than their native Balochi before entering school. Balochi is the language of the home and the surrounding community, in which not only family and friends, but also daycare staff, shopkeepers, and officials speak Balochi. The children may have had limited exposure to Persian via radio and television, where broadcasts are largely in Persian. Regular and extended exposure to Persian did not however occur until primary school. Here, the medium of instruction is Persian, as is the rule in Iran. Teachers are either native speakers of Persian or proficient L2 speakers of Persian (with Balochi as their L1). At primary schools in Balochistan in Iran, Persian is taught for five hours a day for approximately 28 weeks (or 7 months) per year. In first grade, the focus is on learning the basics of the language by means of listening, speaking, and literacy exercises. From second grade, there is an increased focus on the training of reading and writing. Towards the end of the second year, the children are expected to be able to write simple compositions in Persian.

All texts and literacy artifacts the children encounter are in Persian, and they are taught to read and write in Persian only. In class, the children are exclusively exposed to Persian and are encouraged to use Persian in class themselves, whilst Balochi is spoken during breaks and after school. Whilst there may be some variation in the amount of Persian that the children encounter outside school via Persian-language media and friends, we are reasonably confident that the amount of Persian input at school is relatively similar for the children of a particular age in the study.

Written data in the form of a hand-written composition on the topic “Mother” was collected from second- and third-grade pupils at the end of the school year (some in May 2009, some in June 2011). This topic was chosen because it is culturally universal and unlikely to cause problems with school authorities. By that time, the second-graders (age 8–9) had had approximately 1000 hours of exposure to Persian at school. The end of second grade was chosen as this is the earliest point at which the pupils could be expected to be able write a composition on their own. There were 107 pupils in the second-grade group. One of the authors (S. Mohammadi) gathered these data with the help of teachers and administrators at two primary schools. The pupils were given the same topic by their teachers and asked to write a free composition in class with a time limit of about one hour. Care was taken not to put pressure on the pupils and not to present the composition as an exam paper or an assignment that would be graded by the teacher.<sup>13</sup>

The same kind of composition was collected from third-graders (age 10–11), who by then had had ca. 1500 hours of exposure to Persian at school. There were 130 pupils in that group. The pupils were unaware that object marking would be investigated in their texts.

As a control group, 133 monolingual Persian age peers were included in the study. Hand-written compositions on the same topic were collected from third-grad-

<sup>13</sup> We also collected compositions on the topic ‘Prayers’ that was set as an end-of-term exam, but did not include data from these in the study.

ers (age 10–11) with Persian as their L1. The writing task was administered in the same manner as for the L2 pupils. For the L1 control group, care was taken to only include pupils from monolingual Persian homes and to keep the social background of the children as similar as possible to that of the Balochi pupils. Some of the parents of the L1 Persian children were illiterate; others had primary, and some middle school education. As it was easier to find monolingual Persian children outside Balochistan, 55 compositions were collected from pupils attending a primary school in the city centre of Teheran (Arjantin area, Zeynabiyeh Primary) in May 2009. We subsequently deemed it necessary to enlarge this L1 data set, and so 78 additional compositions were collected in 2011 from children from monolingual Persian homes who were attending schools in Iranshahr. As no differences could be detected between the L1 Persian compositions from Teheran and Iranshahr, these were combined to make up a control group of 133 L1 Persian texts.

Table 1 provides an overview of the data. As can be seen in Table 1, group sizes, and thus the number of compositions per group, are not completely even. No more than 107 L2 grade two compositions could be collected and used, since a number of children from non-exclusively Balochi-speaking homes had to be excluded. As is typical for elementary learners, the L2 texts in grade two were shorter on average than the texts in the other two groups.

Participant groups	Compositions	Individual text length	Word total for group
L1 Persian controls, grade 3	N - 133	average 150 words, range 100–200 words	19,950
L2 Persian, grade 2	N - 107	average 80 words, range 50–150 words	8,600
L2 Persian, grade 3	N - 130	average 140 words, range 70–200 words	19,500

**Table 1.** Overview of L1 and L2 data

Copies of the handwritten L1 and L2 texts were analyzed, word-counted, and marked up for indirect and direct objects. These objects were classified into and coded for sub-types depending on grammatical function and morpho-syntactic form, i.e. the use of *be*, *rā*, and learner variants thereof. One of the authors, a speaker of Persian and Balochi, carried out the coding and counting by hand. To increase coding reliability, some data were initially also coded by an experienced Iranian linguist and speaker of Persian and Balochi (C. Jahani). Coding disagreements were resolved through mutual discussion. The data were then entered into Excel spreadsheets, which formed the base for the subsequent analysis.

## 6. Analysis and results

### 6.1. Overall frequencies

All three groups produced more direct objects in their texts than indirect ones, as shown in Table 2. This should not come as a surprise as direct objects are generally more frequent in both speech and writing.

	Indirect objects		Direct objects		Word total
	Total	Per text	Total	Per text	
L1-3 <sup>rd</sup> years N - 133	106	0.8	605	4.5	19,950
L2-2 <sup>nd</sup> years N - 107	186	1.7	405	3.8	8,600
L2-3 <sup>rd</sup> years N - 130	266	2.0	866	6.7	19,500

**Table 2.** Occurrence of indirect and direct objects in the children's compositions

What did come as a surprise however was the fact that the L2 learners produced more indirect objects than the L1 controls. This tendency was pronounced both for the low proficiency learners (grade 2) and the higher proficiency learners (grade 3) who used more than twice as many indirect objects per text than did the native controls (L2ers: 1.7 and 2.0 instances, L1ers: 0.8 instances). The L2 second-graders used direct objects twice as often as indirect objects (405:186), whilst the L1 controls produced six times as many direct objects as indirect objects (605:107). The L2 third-graders produced many more direct objects per text (6.7 instances) than the L1 controls (4.5 instances). Differences in group size or text length cannot account for these ratio differences.

A qualitative analysis of the data showed that some of the L2 learners wrote down the same sentence several times in their text about "Mother", and this often was a sentence in which the mother was the indirect object, such as *be mādar-am ehterām mi-gozār-am* (to mother-my respect IMPF-put-1SG) 'I respect my mother'.<sup>14</sup> This sentence and versions thereof were often used twice and sometimes three or four times in a text by the learners, which increased the number of indirect objects in the L2 groups to a level far beyond that of the L1 controls. We are not quite sure why such repetition of sentences occurred in the L2 texts only. Young school children sometimes have the mistaken impression that a composition is only good or "finished" once the sheet is filled, and it was the youngest learners who repeated sentences the most, even though our task instructions did not invite or urge pupils to do so. Another reason might be that processing demands were simply so high for the L2 learners that consciously or unconsciously relief was sought through the repeated use of the same phrase or sentence, a phenomenon not unknown in second language studies. Note that the L1 pupils did not repeat their own sentences in the texts.

A qualitative analysis of the texts also provided an explanation for the high number of direct objects in the third-grade L2 texts. These often contained short, minimally varied sentences, with the same subject and/or verb repeated again and again, with only the direct object having been changed (e.g. My mother buys a bag for me. My mother buys a dress for me. My mother buys shoes for me. My mother sews a dress for me.). This writing pattern drove up the number of direct objects per text (6.7) as compared to the L1ers (4.5), who did not make much use of such repetitive language.

<sup>14</sup> The Persian for 'I respect my mother' (lit. I put respect to my mother) is:

(i) *man be mādar-e khod ehterām mi-gozār-am*  
I to mother-LINK self respect IMPF-put.PRES-1SG

The L2 Persian learners used this sentence repeatedly, although the second-graders often substituted *rā* for *be*:

(ii) *man mādar-e khod-rā ehterām mi-gozār-am*  
I to mother-LINK self-RA respect IMPF-put.PRES-1SG

In both the L1 and the L2 texts we find sentences with an indirect object only, sentences with a direct object only, and sentences with both indirect and direct object. In the following sections, indirect and direct objects will be investigated separately. We will start with indirect objects.

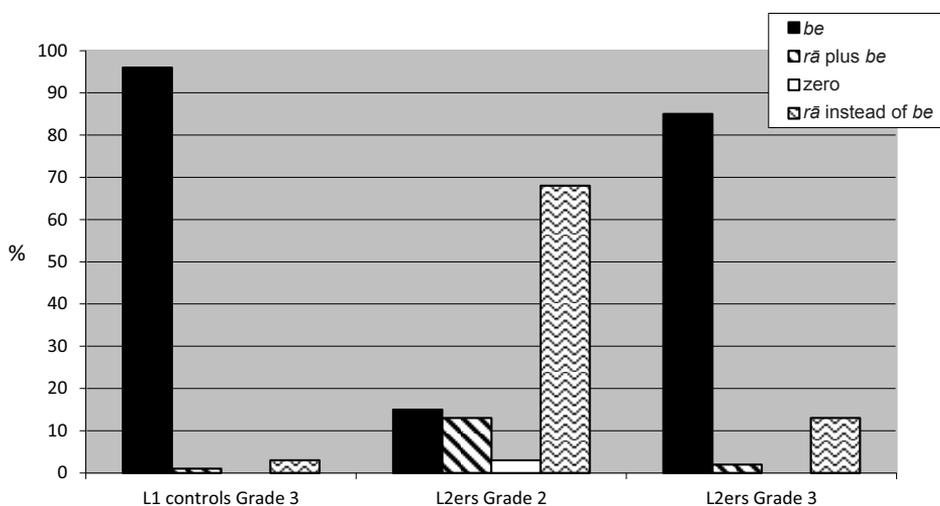
## 6.2. Acquisition of indirect object marking

When looking at the morphological marking of indirect objects, we find large differences between the groups, which are summarized in Table 3 and Fig. 1 and discussed below.

The morphological form of the indirect objects in the 133 L1 Persian pupils' compositions is as would be expected from reference grammars of Persian: They mark virtually all of their indirect objects with the prenominal preposition *be* 'to' (96%, 102/106). Only four indirect objects diverge from this, being non-targetlike uses of postnominal *rā* (i.e. a direct object marker) or both *rā* and *be* on the indirect object. Due to the very low frequency of these non-adultlike structures we may dismiss them as simple writing mistakes, although we cannot entirely rule out that they could also be vestiges of an earlier developmental stage, as we do not have any L1 data from grade two, when these native Persian children started learning to write. Overall, however, the native Persian pupils' indirect object marking with *be* can be considered adultlike.

	<i>be</i>	Simultaneous <i>rā</i> and <i>be</i>	Zero marking	<i>rā</i> instead of <i>be</i>
L1-3 <sup>rd</sup> yrs N - 133	96.2% (102/106)	0.9% (1/106)	0% (0/106)	2.8% (3/106)
L2-2 <sup>nd</sup> yrs N - 107	15.0% (28/186)	13.4% (25/186)	3.2% (6/186)	68.3% (127/186)
L2-3 <sup>rd</sup> yrs N - 130	85.0% (226/266)	1.9% (5/266)	0% (0/266)	13.2% (35/266)

**Table 3.** Indirect object marking in the compositions



**Figure 1.** Indirect object marking in Persian

## 6.2.1. Indirect object marking in the second-grade L2 learners

In stark contrast to the native Persian pupils, the L2 learners in grade two rarely use prepositional *be* to mark indirect objects in their writing (15%, 28/186). The 107 learners only produce 28 instances of adultlike prenominal *be*-marking altogether. Instead, they mostly employ non-targetlike postnominal *rā* (68%, 127/186), as illustrated in examples (30)–(33), where the a-examples give the L2 version and the b-examples the Persian target version.

- (30) a. *mā* *bāyad* *moʔalem-rā* *salām kon-im*  
 we must teacher-RA greeting do.PRES-1PL  
 ‘We must greet the teacher.’ (L2 learner, grade 2, non-target)
- (30) b. *mā* *bāyad* ***be*** *moʔalem* *salām kon-im*  
 we must to teacher greeting do.PRES-1PL  
 (Persian target)
- (31) a. *moʔalem* *mā-rā* *dars mi-āmuz-ad*  
 teacher we-RA lesson IMPF-teach.PRES-3SG  
 ‘The teacher gives/teaches us a lesson.’ (L2 learner, grade 2, non-target)
- (31) b. *moʔalem* ***be*** *mā* *dars mi-āmuz-ad*  
 teacher to we lesson IMPF-teach.PRES-3SG  
 (Persian target)
- (32) a. *man* *moʔalem-e khod-rā* *kādo mi-dah-am*  
 I teacher-LINK self-RA gift IMPF-give.PRES-1SG  
 ‘I give a gift/gifts to my teacher.’ (L2 learner, grade 2, non-target)
- (32) b. *man* ***be*** *moʔalem-e khod* *hedye mi-dah-am*  
 I to teacher-LINK self gift IMPF-give.PRES-1SG  
 (Persian target)
- (33) a. *moʔalem* *bage-hā-rā* *mi-guy-ad*  
 teacher child-PL-RA IMPF-tell.PRES-3SG  
 ‘The teacher tells the children...’ (L2 learner, grade 2, non-target)
- (33) b. *moʔalem* ***be*** *bage-hā* *mi-guy-ad*  
 teacher to child-PL IMPF-tell.PRES-3SG  
 (Persian target)

A further 13% (25/186) of indirect objects produced by the second-grade L2 learners are simultaneously marked with prenominal *be* and postnominal *rā*, as illustrated in (34)–(36). This is not acceptable in native Persian. Simultaneous *be* + *rā* may suggest a beginning awareness of prenominal *be*, at a time when postnominal marking of indirect objects has not yet been expunged from the learners’ interlanguage system.

- (34) a. *be mā-rā mi-guy-ad*  
to we-RA IMPF-tell.PRES-3SG  
'She tells us.' (L2 learner, grade 2, non-target)
- (34) b. *be mā mi-guy-ad*  
to we IMPF-tell.PRES-3SG  
(Persian target)
- (35) a. *khodā be u-rā komak mi-kon-ad*  
God to (s)he-RA help IMPF-do.PRES-3SG  
'God helps her.' (L2 learner, grade 2, non-target)
- (35) b. *khodā be u komak mi-kon-ad*  
God to (s)he help IMPF-do.PRES-3SG  
(Persian target)
- (36) a. *be mā-rā dars mi-dah-ad*  
to we-RA lesson IMPF-give.PRES-3SG  
'She teaches us/gives us a lesson.' (L2 learner, grade 2, non-target)
- (36) b. *be mā dars mi-dah-ad*  
to we lesson IMPF-give.PRES-3SG  
(Persian target)

We thus find that the L2 learners in second grade greatly overuse *rā*: They use it to mark indirect objects, whilst it is restricted to direct objects in Persian.

From a contrastive perspective of Persian and Balochi, these results are to be expected if we take an L1 transfer-driven approach to second language learning. The use of postnominal *rā* on most indirect objects produced by the second-grade learners can be explained as structural and morphological transfer from the L1 to the L2 (e.g. Jarvis and Odlin 2000; Lardiere 2009). Recall that L1 Balochi marks indirect objects with a suffix, not with a preposition. And so the Balochi learners of Persian also latch on to a suffix to mark indirect objects in their L2 Persian. And they do not just latch onto any suffix, but one that is phonologically similar to the one used in L1 Balochi, *-a* (which in some cases allomorphs as *rā*). It is thus not surprising that the learners employ *rā* for indirect object marking in their L2 Persian in analogy with *-a* in their L1. The learners appear to have largely persianised or relexified their Balochi grammar but not yet acquired the target Persian grammar of *be*-marking indirect objects. Consider the (non-target) L2 utterances in (30') and (33') and the corresponding Balochi forms, both with suffixal object marking:

- (30') a. *mā bāyad moʔalem-rā salām kon-im*  
we must teacher-RA greeting do.PRES-1PL  
'We must greet the teacher.' (Persian L2 learner, grade 2, non-target)
- (30') b. *mā bāyad-ē mallem-a salām be-day-ē*  
we must-COP.PRES.3SG teacher-OBJ greeting SBJ-give.PRES-1PL  
(Balochi)

- (33') a. *moʔalem*    *bache-hā-rā*    *mi-guy-ad*  
 teacher    child-PL-RA    IMPF-tell.PRES-3SG  
 'The teacher tells the children.' (Persian L2 learner, grade 2, non-target)
- (33') b. *malle*    *chokk-ān-a*    *gwash-i*  
 teacher    child-PL-OBJ    tell.PRES-3SG  
 (Balochi)

The L2ers' *rā* on indirect objects may also be regarded as an overextended default marking. The second-graders make ample use of *rā* for direct objects (see section 6.3). Direct objects are more frequent than indirect ones, and overgeneralizing *rā* to the latter would not be an unusual acquisitional route to take. It is likely, though, that such overextension is not so much motivated by the system of Persian, but by the L1 system of Balochi where there is one morpheme (-*a*) for both direct and indirect object marking. Under a frequency-driven approach to second language acquisition, the learner structure may be attributed to *rā* being a frequent morpheme in Persian, which presumably is frequent in the input as well, although we currently lack frequency counts for Persian corpora in general and for teacher talk (the main source of Persian input for our learners) in particular. *Rā* is invariant in form and word-final, which may contribute to its salience, and it occurs on direct objects (which are more frequent than indirect ones) as well as aboutness topics (see fn. 6).

At the same time, we can also see that the L2ers have some modest awareness of *be*-marking in Persian. Recall Table 3, and the 28 instances (15%) of targetlike *be*-marking, produced by 19 of the 107 learners. (Six of these learners also produce non-targetlike Balochi-style *rā* on indirect objects.) Recall also the 25 instances (13%) of simultaneous *be* and *rā*, produced by 17 of the 107 learners (as exemplified in (33)–(35) above). We thus see the beginnings of a development towards the Persian target structure.

Interestingly, there is hardly any zero marking of indirect objects in the learner data (3%, 6/186 instances); that is, pronouns and nouns used as indirect objects do not occur bare but nearly always with some (albeit non-targetlike) object marker. In studies of second language production, bare and uninflected forms are usually frequent, especially in the beginning stages of acquisition, and functional morphology (such as case marking) is often found to be missing. By contrast, our beginning learners of Persian do not just omit *be* but actually use overt suffixal marking on indirect objects, which suggests that they indeed transfer their L1 Balochi morphological object marking to their L2 Persian.

### 6.2.2. Indirect object marking in the third-grade L2 learners

Turning to the more advanced learners, the pupils in third grade (Table 3), we can see that they mark indirect objects most of the time with prepositional *be* (85%, 226/266). This percentage is very different from the second-grade L2ers (15%), but relatively close to the provision of *be* by the age-matched native L1 controls (96%). Some of the L2 third-graders still produce Balochi-style *rā* instead of *be* (13%, 35/266). This is done by 25 out of 130 learners, and is exemplified in (37)–(38). There

are however very few children in this group that still only use Balochi-style *rā* instead of *be*; nearly all the L2 third-graders (125 out of 130 children) use targetlike indirect object marking with prepositional *be* at least once.

- (37) a. *mādar-e khod-rā*                      *komak mi-kon-am*  
 mother-LINK self-RA                      help IMPF-do.PRES-1SG  
 ‘I help my mother.’ (L2 learner, grade 3, non-target)
- (37) b. *be mādar-e khod*                      *komak mi-kon-am*  
 to mother-LINK self                      help IMPF-do.PRES-1SG  
 (Persian target)
- (38) a. *moʔalem*                      *mā-rā*                      *khāndan o neveshtan āmukht*  
 teacher                      we-RA                      reading and writing teach.PAST.3SG  
 ‘The teacher taught us how to read and write.’  
 (L2 learner, grade 3, non-target)
- (38) b. *moʔalem*                      *be mā*                      *khāndan o neveshtan āmukht*  
 teacher                      to we                      reading and writing teach.PAST.3SG  
 (Persian target)

No cases of zero marking occur; there are however still a few cases of simultaneous *be* and *rā* on indirect objects (2%, 5/266) in the third-grade L2ers. Apart from these non-target forms, the third-graders have largely mastered Persian indirect object marking, as far as can be determined from their compositions.

We do not have the longitudinal data for the same individuals needed to say for sure that they indeed go from non-target *rā* marking (and optionally via interim simultaneous *be* and *rā*) to targetlike exclusive *be*-marking. Nevertheless, the comparison of the two learner groups (grades 2 and 3) suggests a clear developmental path towards the native Persian pattern.

### 6.3. Acquisition of direct object marking

All three groups of school children produce many instances of direct objects in their compositions (recall Table 2). When looking at the morphological marking of direct objects, we do not find quite as large differences between the groups as we did for indirect objects, but the differences between groups are still sizeable. These concern the different morphological forms, as well as extent to which *rā* marking is overt. To start with, let’s look at the different forms of morphological marking in the data, summarized in Table 4 and Figure 2.

	<i>rā</i>	zero marking	multiple <i>rā</i>	<i>be</i> instead of <i>rā</i>
L1-3 <sup>rd</sup> years	75.5% (457/605)	23.8% (144/605)	0.5% (3/605)	0.2% (1/605)
L2-2 <sup>nd</sup> years	34.6% (140/405)	49.6 % (201/405)	10.1% (41/405)	5.7% (23/405)
L2-3 <sup>rd</sup> years	65.7% (569/866)	33.3% (288/866)	1.0% (9/866)	0.0% (0/866)

**Table 4.** Direct object marking in the compositions

The L1 Persian pupils in third grade mark a large majority of their direct objects (76%, 457/605) with the postnominal *rā*. The group is relatively homogeneous here;

most L1 children produce such overt *rā*. The remaining 24% of their objects are not suffixed (“zero marking”), an option Persian has due to DOM. We may assume that this distribution of 76% overt *rā* and 24% zero marking on direct objects is a typical distribution for Persian and may serve as a benchmark, at least for the type of written texts that we are dealing with here. Only 0.7% of the L1ers’ direct objects are morphologically non-adultlike: There are 3 instances of multiple *rā* (i.e. more than one *rā* on the same object), and 1 prepositional *be* instead of *rā*. Considering the extremely low frequency of these cases, they are most likely to be simple writing mistakes.

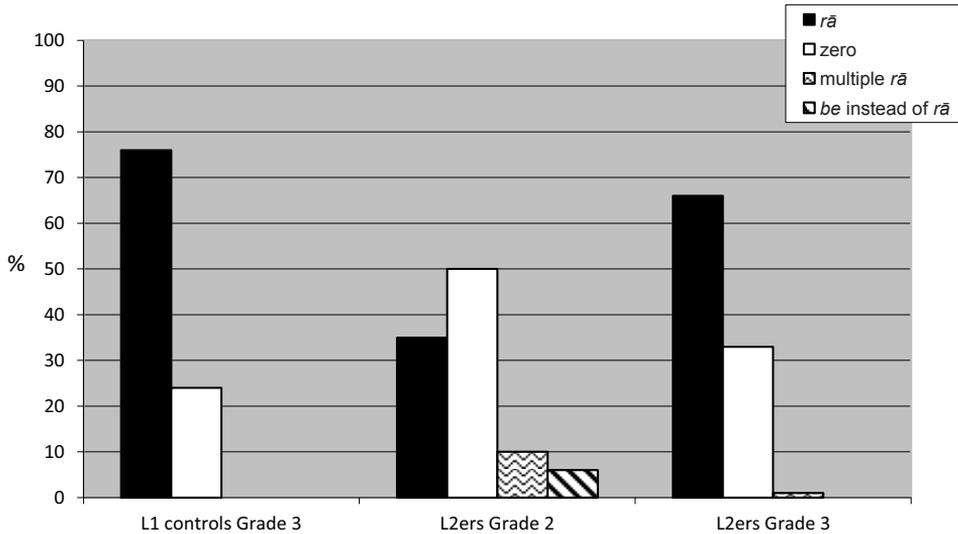


Figure 2. Direct object marking in Persian

### 6.3.1. Direct object marking in the second-grade L2 learners

The L2 data are quite different from the native controls (Table 4). Only a third of the second-grade L2ers’ direct objects are overtly marked with *rā* (35%, 140/405); the majority of their direct objects are zero marked (50%, 201/405). Whilst zero marking certainly is an option in Persian due to DOM, it is nevertheless striking that the learners make so much more use of zero marking than the L1 controls. We will get back to these zero markings in the discussion below (section 6.3.3).

The L2ers’ 35% *rā* on direct objects is still a sizeable percentage, and the group is relatively homogeneous here: Most of the second-graders (83%, 89 out of 107) produce *rā* on direct objects at least once. This is quite different from their indirect object marking, where only 18% (19 out of 107 children) produced an instance of targetlike indirect object marking (recall section 6.2.1).

Apart from 35% *rā* and 50% zero marking, the second-grade L2ers also produce some deviant overt direct object markings: 10% (41/405) of their direct objects are marked twice or three times with *rā*, as in (39)–(41). This is not an option in native Persian. Whilst one might want to discount the L2ers’ multiple *rā* cases as simple writing mistakes, it is nevertheless noteworthy that so many mistakes of this kind

occur in the L2 data (10%), compared to only 0.5% in the L1ers. 34 out of 107 second-grade L2 children produce multiple *rā* at least once.<sup>15</sup>

(39) a. *moʔalem*      *dars-rā-rā*      *mi-dah-ad*  
 teacher      lesson-RA-RA      IMPF-give.PRES-3SG  
 ‘The teacher teaches the lesson.’ (L2 learner, non-target multiple *rā*)

(39) b. *moʔalem*      *dars-rā*      *mi-dah-ad*  
 teacher      lesson-RA      IMPF-give.PRES-3SG  
 (Persian target)

(40) a. *mā na-bāyad*      *u-rā u-rā*      *aziyat kon-im*  
 we NEG-must      (s)he-RA (s)he-RA      annoy do.PRES-1PL  
 ‘We must not annoy her.’ (L2 learner, non-target multiple *rā*)

(40) b. *mā nā-bāyad*      *u-rā*      *aziyat kon-im*  
 we NEG-must      (s)he-RA      annoy do.PRES-1PL  
 (Persian target)

(41) a. *man*      *u-rā kheyli-rā*      *dust dār-am*  
 I      (s)he-OBJ much-RA      love have.PRES-1SG  
 ‘I love her very much.’ (L2 learner, non-target multiple *rā*)

(41) b. *man*      *u-rā kheyli*      *dust dār-am*  
 I      (s)he-OBJ much      love have.PRES-1SG  
 (Persian target)

Another non-target form that the second-grade L2ers use is prepositional *be* instead of suffixal *rā* (6%, 23/405), as in (42). Only a minority of children (17 out of 107) use this form.

(42) a. *mādar-am*      *be man*      *ne-mi-zan-ad*  
 mother-my      to I      NEG-IMPF-hit.PRES-3SG  
 ‘My mother does not beat me.’ (L2 learner, non-target *be* instead of *rā*)

(42) b. *mādar-am*      *man-rā*      *ne-mi-zan-ad*  
 mother-my      I-OBJ      NEG-IMPF-hit.PRES-3SG  
 (Persian target)

The use of *be* instead of *rā* suggests that a few second-grade children have noticed the occurrence of *be* in connection with objects in Persian in their input and are beginning to use the form themselves. They appear to be doing so, however, without

<sup>15</sup> Sometimes there is double *rā* on the same noun, and sometimes *rā* appears twice on two adjacent words. We have at present no good explanation for this high occurrence of non-target multiple *rā* in the L2 second-graders. Could processing play a role here? These pupils are the youngest and least experienced/fluent in writing. Presumably they devote proportionally more of their processing capacity to transcription and orthographic coding than the third-grade groups need to. Temporary cognitive overload might therefore lead to more repetitions and multiple *rā*. In order to shed light on this, one should compare the second-grade L2ers with a group of age-matched L1 native Persian pupils who have as little experience in writing as they do. We leave this matter for future research.

having fully understood the Persian pattern, where *be* is restricted to indirect objects and *rā* is restricted to direct objects. As the children's L1 Balochi does not make a formal distinction between indirect and direct object marking, it is to be expected that it will take the learners some time to figure out the distributional pattern of Persian and implement it in their own L2 Persian.

The relatively low use of *rā* in general, coupled with unwarranted *rā* doubling on some direct objects, and the non-targetlike use of *be* with direct objects suggests that there is considerable confusion about Persian object marking in the second-grade L2 learners at this stage. Recall also that *be* marking on indirect objects only rarely occurs at this stage. It thus appears that a few learners have noticed the form *be* and added it to their lexicons but are not yet able to use it in a targetlike fashion, resulting in *be* overuse with direct objects and severe underuse with indirect objects.

### 6.3.2. Direct object marking in the third-grade L2 learners

The picture is quite different in the third-grade L2 compositions. Here, a majority of direct objects, 66% (569/866), are marked with overt *rā*, not unlike the L1ers (who, as will be recalled, used 76% overt *rā*). Nearly everyone in the group (125 out of 130 children) produces such overt *rā*. Only very few of the L2 third-graders' direct objects contain non-target multiple *rā* (1%, 9/866, produced by 7 out of 130 children).

The other non-target option that occurred in the second-grade learners, namely using prepositional *be* instead of suffix *rā*, is not attested in grade three. The L2 third-graders are thus fairly similar to their age-matched native Persian controls. Even though we do not have longitudinal data for individual learners, the comparison of the second-grade and third-grade learner groups indicates a clear development towards target Persian.

### 6.3.3. Zero marking on direct objects: DOM revisited

As shown in Table 4 and Fig. 2, there is a considerable difference between the groups concerning the overtness of *rā*: 76% of all direct objects produced by the L1ers show overt *rā*, but only 35% for the L2 second-graders,<sup>16</sup> going up to 66% for the L2 third-graders. This warrants further investigation.

Recall from our discussion of DOM (section 2.2), that *rā* in adult Persian is mostly overt when the direct object is human and/or when it is specific and identifiable from context. By contrast, inanimate and unspecific or generic direct objects are only very rarely marked with *rā* in adult Persian. Our native L1 Persian third-graders fully adhere to this pattern. They use *rā* on animate direct objects with specific reference, for instance ones expressed by personal pronouns e.g. *to-rā* 'you', as in (43), as well as on inanimate direct objects with specific reference, e.g. noun phrases with a possessive determiner, e.g. 'my foodstuffs' (44), 'our clothes' (45). By contrast, the native Persian children do not use *rā* on inanimate and unспе-

<sup>16</sup> Even if we add the second-grade L2ers' non-target multiple *rā* (10%) to their single *rā* (35%), their proportion of overt *rā* marking on direct objects (45%) falls substantially below that of the native Persian children (76%).

cific or generic direct objects, e.g. ‘good things’, ‘lunch’, or ‘delicious food’, as illustrated by the authentic examples in (46)–(48).

Overt *rā* on direct objects with specific reference:

- (43) *mādar-am*      *man*      *to-rā*      *dust dār-am*  
 mother-my      I      you-OBJ      love have.PRES-1SG  
 ‘Oh my mother, I love you.’ (L1 third grade)
- (44) *mādar-am*      *khōrāki-hā-yam-rā*      *dar kif-am*      *mi-gozār-ad*  
 mother-my      foodstuffs-PL-my-OBJ      in bag-my      IMPF-put.PRES-3SG  
 ‘My mother puts my foodstuffs in my bag.’ (L1 third grade)
- (45) *mādar*      *lebās-hā-ye mā-rā*      *tamiz*      *mi-kon-ad*  
 mother      clothes-PL-LINK our-OBJ      clean      IMPF-do.PRES-3SG  
 ‘Mother cleans our clothes.’ (L1 third grade)

Zero marking (no overt *rā*) on direct objects with unspecific or generic reference:

- (46) *mādar-am*      *barā-ye mā*      *chiz-hā-ye khub*      *mi-khar-ad*  
 mother-my      for-LINK we      thing-PL-LINK good      IMPF-buy.PRES-3SG  
 ‘My mother buys good (i.e. beautiful) things for us.’ (unspecific) (L1 third grade)
- (47) *vaghti*      *az madrese*      *mi-āy-am*  
 when      from school      IMPF-come.PRES-1SG  
    *mādar-am*      *nāhār*      *āmāde mi-kon-ad*  
    mother-my      lunch      prepare IMPF-do.PRES-3SG  
 ‘When I come back from school, my mother prepares lunch.’ (unspecific)  
 (L1 third grade)
- (48) *mādar-am*      *barā-ye man*      *ghazā-ye khoshmāze*      *mi-paz-ad*  
 mother-my      for-LINK I      food-LINK delicious      IMPF-cook.PRES-3SG  
 ‘My mother cooks delicious food for me.’ (unspecific) (L1 third grade)

For the L1ers then, the distribution of overt *rā* vs. zero marking is 76% vs. 24% (Table 4), and when looking at their direct objects in context, this distribution appears to be adultlike. We have not been able to find any ungrammatical examples with *rā* in the L1 Persian compositions, and only very few instances where an overt *rā* would have been preferable to the zero marking used by the L1 children. This is very different from the L2 learners.

The L2ers, especially in second grade, produce many more zero-marked direct objects (50%) than the L1 Persian children do (24%). In theory, it could be that the L2ers write about many more inanimate and/or unspecific objects than the L1ers, which would explain the difference in overt and zero marking between the two groups. However, an analysis of the L2ers’ productions in context indicates that this is not the case. Rather, they also write about animate objects and objects with specific reference just as the native L1 controls do.

The L2 learners zero-mark many direct objects high up on the animacy hierarchy (typically referring to humans) that have specific reference in context. Such objects would usually occur with overt *rā* in native Persian, and we will therefore refer to

them as *rā* omissions. Some examples are given in (49)–(55), where *\_* indicates the missing *rā*. As can be seen, the learners omit *rā* both from full noun phrase objects and from pronominal objects (e.g. *u-rā* ‘her/him’, *madare khod-rā* ‘my mother’, *man-rā* ‘me’, *bacheha-rā* ‘the children’, *moaleme khod-rā* ‘my teacher’).

- (49) a. *man u\_ kheyli dust dār-am*  
I (s)he much love have.PRES-1SG  
‘I love her/him very much.’ (L2 learner, grade 2, nontarget)
- (49) b. *man u-rā kheyli dust dār-am*  
(Persian target)
- (50) a. *man mādar-e khod\_ dust dār-am*  
I (s)he-LINK self love have.PRES-1SG  
‘I love my mother.’ (L2 learner, grade 2, nontarget)
- (50) b. *man mādar-am-rā dust dār-am*  
(Persian target)
- (51) a. *pedar-am man\_ be Chābahār bord*  
father-my I to Chabahar took.PAST.3SG  
‘My father took me to Chabahar.’ (L2 learner, grade 2, nontarget)
- (51) b. *pedar-am man-rā be Chābahār bord*  
(Persian target)
- (52) a. *mā na-bāyad bache-hā\_ ayizat kon-im*  
we NEG-must child-PL annoy do.PRES-1PL  
‘We must not bother the children (i.e. the other students at school).’  
(L2 learner, grade 2, nontarget)
- (52) b. *mā na-bāyad bache-hā-rā ayizat kon-im*  
(Persian target)
- (53) a. *har vaght mo?alem-e khod\_ did-am ...*  
every time teacher-LINK self saw.PAST-1SG  
‘When I saw my teacher ...’ (L2 learner, grade 2, non-target)
- (53) b. *har vaght mo?alem-e khod-rā did-am ...*  
(Persian target)
- (54) a. *khodā u\_ dust na-dār-ad*  
God (s)he love NEG-have.PRES-3SG  
‘God does not love her/him.’ (L2 learner, grade 3, non-target)
- (54) b. *khodā u-rā dust na-dār-ad*  
(Persian target)
- (55) a. *mā bāyad ghadr-e mo?alem-e khod\_ be-dān-im*  
we must appreciate-LINK teacher-LINK self SBJ-know.PRES-1PL  
‘We must appreciate our teachers.’ (L2 learner, grade 3, non-target)
- (55) b. *mā bāyad ghadr-e mo?alem-e khod-rā be-dan-im*  
(Persian target)

As shown in Table 5, non-target *rā* omissions are frequent in the beginner L2 group (grade 2), who leave out *rā* in a non-targetlike manner 30% of the time (120/405), compared to 1% in the L1 group (7/605). The learners thus underuse *rā*. Whilst not every second-grade L2 child omits *rā*, the majority of them do so some of the time (58 out of 107 children). Sentences of the type illustrated in (49), where *rā* has been omitted from a pronominal direct object, are very frequent in the L2 compositions.

	Adultlike <i>rā</i>	Adultlike zero marking	Non-adultlike zero marking (- <i>rā</i> omission)
L1-3 <sup>rd</sup> years	75.5% (457/605)	22.6% (137/605)	1.2% (7/605)
L2-2 <sup>nd</sup> years	34.5% (140/405)	19.7% (80/405)	<b>29.8% (121/405)</b>
L2-3 <sup>rd</sup> years	65.7% (569/866)	29.4% (255/866)	3.8% (44/866)

**Table 5.** Zero marking on direct objects

The third-grade L2ers also occasionally omit *rā* from direct objects (4%, 44/866), 20 out of 130 children do so. 4% is higher than the native controls (1%), but much less frequent than the second-grade L2ers (30%). Apart from these few non-target forms, the third-graders appear to have mastered Persian direct object marking. Whilst we do not have longitudinal data for the same individuals over time, a comparison of the two learner groups suggests a developmental path towards the native Persian distribution.

Readers may wonder why the 130 third-grade L2ers produce so many more direct objects in their texts (866) than the 133 native L1 controls in theirs (605), and why the L2ers produce a higher proportion of acceptable zero-marked direct objects (29%) than the native L1 controls (23%). An inspection of the compositions suggests that these differences in frequency (but not in grammaticality) are due to slightly different ways of writing (recall section 6.1): The third-grade L2ers often write short sentences with minimal variation, repeating the same subject and/or verb again and again, and only changing the direct object or the verb. The direct object here usually is an unspecific inanimate noun phrase that does not require *rā* (DOM, recall section 2.2).<sup>17</sup> This is grammatically correct but textually repetitive. The L1ers produce such repetitive writing to a much lower degree.

But the greatest difference between the L2 learners and native controls does not concern the third-graders, but the second-graders. As mentioned above, these often omit *rā* from direct objects in a *non-targetlike* manner (Table 5). How can we explain these learner results? Some answers are suggested in the next section.

#### 6.3.4. Discussion of direct object marking results

From a contrastive perspective, it is somewhat surprising that the beginning L2 learners (grade 2) diverge so much from the native Persian controls, since Balochi and Persian are relatively similar in their marking of direct objects. Recall from sections 2 and 3 that both languages mark direct objects with suffixes that are mor-

<sup>17</sup> Corresponding examples in English would be: ‘My mother buys a bag for me. My mother buys a dress for me. My mother buys shoes for me. My mother sews a dress for me.’

pho-phonologically quite similar. Both languages have DOM, where objects high on the animacy hierarchy/scale (especially humans) and objects that are easily context-identifiable (specific reference) are typically marked overtly, whilst inanimate objects with unspecific reference are not.

Now, if learners simply were to transfer their Balochi grammar to Persian (relexification), we would expect to see more of the Persian pattern already in the elementary learners, since Persian direct object marking does not much differ from Balochi. By contrast, indirect object marking in the two languages is not similar, and here transfer results in 82% non-target Balochi-style and Balochi-influenced constructions in the learners (section 6.2).

What about the direct object marking in the elementary learners? Several possibilities spring to mind. Firstly, one could appeal to universal tendencies in second language acquisition to explain the learners' underuse of *rā* marking on direct objects. It is well known that learners, especially at the beginning stages, often omit functional/grammatical markers. It is also well known that learners exhibit morphological variability and optionality in their use of nominal and verbal inflection and associated lexical items. Thus, during the course of acquisition, there is a long stretch of time during which functional markers are sometimes present and sometimes absent in circumstances where they would be (more or less) obligatory for native speakers. So this might also hold for Balochi learners of Persian. Still it is noteworthy that the very same learners who optionally omit *rā* from direct objects make extensive use of all manner of functional morphology in their productions. They produce 45% *rā* on direct objects, and 82% *rā* in another context, namely on indirect objects (where Persian requires *be*, and *be* only). Thus we cannot explain the high number of zero marked direct objects as a lack of functional morphology in their interlanguage grammar or a lack of the functional morpheme *rā* in particular.

One might want to attribute the optional use of *rā* on direct objects to processing problems of some sort. Morphology must be learnt, and the L2ers have clearly learnt the form *rā*. Then why do they not use it consistently? The acquisition literature speaks of a mapping problem, a temporary breakdown between the syntax and the lexicon (e.g. Prévost and White 2000). In the context of text writing, one might further suppose that novice writers like our second-grade learners are prone to cognitive overload, under the strain of L2 sentence formulation, text planning, and handwriting, and sometimes forget to write down *rā* at the end of noun phrases. Whilst this kind of breakdown cannot be ruled out, we would expect the learners to also omit other elements (morphemes, words) in their writing. As far as we can tell, there is no such pattern.

It has been suggested to us (C. Jahani, personal communication) that the learners' omission of direct object marking might be explained via the existence of DOM. Both Balochi and Persian have obligatory overt marking of indirect objects, and transferring this requirement from the L1 to the L2 would result in constant provision of indirect object marking, which is indeed what we find in the learners, albeit largely in non-target form (*-rā* instead of *be*). By contrast, Balochi and Persian do not have obligatory but differential marking of direct objects, and transferring this 'non-obligatoriness' from the L1 to the L2 would result in optional provision of di-

rect object marking. However, if it is true that DOM in the two languages is governed by similar semantic and discourse-pragmatic constraints (sections 2.2 and 3), it remains unclear why Balochi learners of Persian would sometimes omit *rā* on Persian direct objects where *-a* is obligatory in Balochi, such as on context-identifiable animate/human objects.

Alternatively, one could entertain the possibility that Balochi and Persian are not so similar in their marking of direct objects after all, and that subtle contrastive differences between the two languages lie at the heart of the learners' *rā* omissions. Recall that in Balochi, overt object marking (*-a*) can only occur in constructions in the nominative-accusative system (present tense), but not in the ergative system (mainly past tense). On an L1-transfer approach, this split might carry over to the productions of Balochi beginning learners of L2 Persian. There might thus arise an interaction between tense and overt object marking in L2 Persian, where *rā* would be omitted more often in a clause in the past tense than in a sentence in the present tense. Why? Because the L1 would have overt direct object marking in the present tense, but not in the past tense.

We have explored this possibility post-hoc in our data, but the results are inconclusive. As the compositions were written on the topic "Mother", the children predominantly wrote in the present tense, describing their mother, her characteristics and activities, their own school-day, and their return home, etc. For this reason, the compositions only contain a few clauses in the past tense. In those that exist, we do find *rā* omissions (e.g. (51) 'My father took me to Chabahar.', (53) 'When I saw my teacher ...'). There does appear to be a higher proportion of *rā* omissions in L2 past tense clauses than in L2 present tense clauses, but the small number of past tense clauses overall precludes any meaningful statistical investigation.

The great majority of clauses in the compositions are in the present tense, for native Persian pupils and L2 learners alike. In these present tense clauses, we also find *rā* omissions for the learners (e.g. (49), (50), (52)), but in other present tense clauses, *rā* is provided, or a non-target form (multiple *rā* or *be* instead of *rā*) is used. In short, we have not been able to find a straightforward interaction of tense and overtness of *rā* marking in our learner data, but then again, the compositions are not ideal for testing this possibility.<sup>18</sup> We aim to investigate this issue further in future work. A new learner corpus is currently being collected, with compositions on the topic "How did you spend your last summer?", which is intended to induce pupils to write predominantly in the past tense. We hope that a comparison of these compositions with the current material (predominantly in the present tense) will enable us to do more justice to the question of tense and overtness of *rā* marking.

<sup>18</sup> Some expressions in Balochi can be constructed ergatively irrespective of tense (e.g. the verb *dust* 'love', recall footnote 8), and the direct object in these constructions is not overtly marked. If we extend the idea of L1 transfer from these constructions to L2 Persian, we would expect that the learners should not use any overt direct object marker when constructing the corresponding sentences in Persian, which would result in non-target *rā* omissions. It is indeed the case that the second-grade learners often omit *rā* from sentences with *dust* 'love' (see e.g. examples (49)–(50), and the third-grade learners sometimes do so too (e.g. (54)). Overall however, there are too few examples to draw any strong conclusions.

## 7. Conclusion

In this paper, we have investigated the acquisition of object marking in the L2 Persian writing of Balochi-speaking children aged eight to eleven years. 107 elementary learners in second grade (after 1000 hours of exposure to Persian), 130 more proficient learners in third grade (after 1500 hours of exposure), as well as 133 native Persian L1 children of similar socio-economic background wrote a composition on a set topic. In these compositions, an analysis of the frequencies and types of indirect and direct object marking revealed large differences between the groups. The native Persian children basically performed according to adult norms of object marking (for 96% of the indirect objects and 98% of the direct objects).

This was in stark contrast to the second-grade L2ers, who predominantly employed non-target forms of object marking in their Persian, but which closely resembled Balochi constructions: 68% of their indirect objects were not marked with prepositional *be* but with suffixal *rā* (resembling Balochi *-a*). Nearly every second-grade learner used this non-target form. Only in a minority of cases (15%) did the learners use targetlike prepositional *be* as the only object marker, whilst some created novel combinations of simultaneous *be* and *rā*. We suggested that these findings could be interpreted as initial transfer of the L1 Balochi suffixal object marking construction to the interlanguage grammar, with subsequent incipient acquisition of the target prepositional construction by a few learners. The more advanced third-grade L2ers used prepositional *be* marking on 85% of their indirect objects. This was interpreted as the learners having largely rid themselves of Balochi-style suffixal marking and having acquired the Persian target form at this stage, at least as could be determined from their L2 writing.

Concerning direct object marking, the second-grade L2ers again differed starkly from the native controls: Whilst the native Persian children mostly used overt *rā* marking (76%) on their direct objects, the L2 learners only did so 35% of the time. The finding for native Persian is interesting, because it, for the first time as far as we are aware, provides figures for differential object marking (DOM), i.e. for the frequency distribution of overt direct object marking in Persian: 76% overt vs. 24% zero, at least as can be ascertained from the ca. 20,000 words of the native written genre sampled here. The elementary L2 learners exhibited a different pattern: Whilst L1ers and L2ers used the same form they did so to very different degrees. This was surprising, as Persian and Balochi are usually said to have the same type of differential object marking (DOM) on direct objects: suffixal *rā/-a*, its overtness being constrained by semantic and discourse-pragmatic factors, with human/animate and specific/context-identifiable objects marked overtly, and inanimate, unspecific objects typically remaining bare. On an L1-transfer approach to acquisition, one would predict few learning problems in an area where the two languages are strictly parallel. However, our learners presented a different picture: Elementary second-grade L2ers often omitted direct object marking in a non-targetlike fashion (30%) or, less commonly, created novel ways of marking objects, different from their L1 Balochi and different from native Persian. The more advanced third-grade L2ers, however, used direct object marking in a fashion

much like their native Persian peers. We discussed a number of alternative explanations for the surprising omissions of object marking in the elementary learners and suggested that DOM in Balochi and Persian might not be as parallel after all. Due to its split ergativity system, Balochi only marks direct objects in certain tenses, which reduces the frequency of overt object marking compared to Persian (which does not have split ergativity). As our learner data were mainly in the present tense, we could not fully test the explanatory strength of this tense/object marking interaction, but hope to do so in future work.

Overall, we could confirm that object marking, as has been suggested in the literature, may be a problematic area for L2 learners of Persian, but we could also see a major improvement between the two learner groups with just one extra year of schooling. However, all our data concern written Persian. It remains to be seen how quickly Persian object marking is acquired and mastered by second language learners in the spoken modality, and how learners would perform on grammaticality or acceptability tests that involve object marking.

## Abbreviations

1	first person	NEG	negation
2	second person	NOM	nominative
3	third person	OBL	oblique
Bal.	Balochi	OBJ	object
CL	clitic	PAST	past
COP	copula	Per.	Persian
DEF	definite	PL	plural
DOM	differential object marking	PRES	present
IMPF	imperfective	SBJ	subjunctive
INDEF	indefinite	SG	singular
LINK	linking particle (ezāfe)		

## References

- Bosson, Georg (1985). *Empirische Universalienforschung: Differentielle Objektmarkierung in den neu-iranischen Sprachen*. Tübingen: Gunter Narr.
- Bowles, Melissa, and Montrul, Silvina (2009). "Instructed L2 acquisition of Differential Object Marking in Spanish." In: Leow, Ronald, Campos, Héctor, and Lardiere, Donna (eds), *Little Words: Their History, Phonology, Syntax, Semantics, Pragmatics and Acquisition*. Washington, DC: Georgetown University Press, pp. 199–210.
- Browne, Wayles (1970). "More on Definiteness Markers: Interrogatives in Persian." *Linguistic Inquiry* 1, pp. 359–363.
- Comrie, Bernard (1989). *Language Universals and Linguistic Typology*. 2<sup>nd</sup> ed. Oxford: Blackwell.
- Dabir-Moghaddam, Mohammad (1990). Pirāmun-e rā dar zabān-e fārsi [On postposition rā in Persian]. *Iranian Journal of Linguistics* 7, pp. 2–60.
- Farrell, Tim (1990). *Basic Balochi: An Introductory Course*. Naples: Istituto Universitario Orientale.
- Foroodi-Nejad, Farzaneh (2011). *Towards the Identification of Linguistic Characteristics of Specific Language Impairment in Persian*. PhD thesis. Edmonton: University of Alberta.
- Ganjavi, Shadi (2007). *Direct Objects in Persian*. PhD thesis. University of Southern California.
- Ghomeshi, Jila (1996). *Projection and Inflection: A Study of Persian Phrase Structure*. PhD thesis, University of Toronto.

- Ghomeshi, Jila (1997). "Topics in Persian VPs." *Lingua* 102, pp. 133–167.
- Ghomeshi, Jila (2003). "Plural Marking, Indefiniteness and the Noun Phrase." *Studia Linguistica* 52, pp. 47–74.
- Jahani, Carina (2003). "The Case System in Iranian Balochi in a Contact Linguistic Perspective." In: Jahani, Carina and Korn, Agnes (eds) *The Baloch and their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*. Wiesbaden: Reichert, pp. 113–132.
- Jahani, Carina and Korn, Agnes (2009). "Balochi." In: Windfuhr, Gernot (ed.), *The Iranian Languages*. London and New York: Routledge, pp. 634–694.
- Jarvis, Scott and Odlin, Terence (2000). "Morphological Type, Spatial Reference, and Language Transfer." *Studies in Second Language Acquisition* 22(4), pp. 525–556.
- Karimi, Simin (1989). *Aspects of Persian Syntax, Specificity, and the Theory of Grammar*. PhD thesis, University of Washington.
- Karimi, Simin (1990). "Obliqueness, Specificity, and Discourse Functions: *rā* in Persian." *Linguistic Analysis* 20(3–4), pp. 139–191.
- Karimi, Simin (1996). "Case and Specificity. Persian *rā* Revisited." *Linguistic Analysis* 26, pp. 174–194.
- Karimi, Simin (2003). "Object Positions, Specificity and Scrambling." In: Karimi, Simin (ed.), *Word Order and Scrambling*. Oxford: Blackwell, pp. 91–125.
- Keshavarz, Mohammad Hossein (2007). "Morphological Development in the Speech of a Persian-English Bilingual Child." *Journal of Psycholinguistic Research* 36, pp. 255–272.
- Korn, Agnes (2009). "The Ergative System in Balochi from a Typological Perspective." *Iranian Journal of Applied Language Studies* 1(1), pp. 43–79.
- Lardiere, Donna (2009). "The Contrastive Analysis of Features in Second Language Acquisition." *Second Language Research* 25(2), pp. 173–227.
- Lazard, Gilbert (1982). "Le morphème *rā* en Persan et les relations actanciennes." *Bulletin de la société de linguistique de Paris* 73(1), pp. 177–208.
- Lazard, Gilbert (1984). "Actance Variations and Categories of the Object." In: Plank, Frans (ed.), *Objects: Towards a Theory of Grammatical Relations*. London: Academic Press, 269–292.
- Lazard, Gilbert (1992). *A Grammar of Contemporary Persian*. Mazda: Costa Mesa, CA and New York. [English translation of *Grammaire du persan contemporain* 1957].
- Lyons, Christopher (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Mirza, Zahra (2000). *Yādgiri-ye maqulehā-ye kār-kardi dar zabānhaye fārsi va armani tavasot-e faragirān-e an zabānhā* [The acquisition of functional categories in Persian and Armenian by learners of those languages] MA thesis. Esfahan: Esfahan University.
- Mohammadi, Somayeh (2009). *Tahlil-e khatāhā-ye nahvi-ye dāneshāmuzān-e Baluchzabān dar yādgiri-ye zabān-e fārsi* [The analysis of syntactic errors of Baloch pupils learning Persian]. MA thesis. Tehran: Allāme Tabātabā'i University.
- Prévost, Philippe and White, Lydia (2000). "Missing Surface Inflection or Impairment in Second Language Acquisition? Evidence from Tense and Agreement." *Second Language Research* 16, pp. 103–134.
- Rahimian, Jalal and Hajiani, Farrokh (2009). "Semantic-Pragmatic Function of *rā* in Persian: A Diachronic and Synchronic Study." *Journal of Indo-European Studies* 37(3–4), pp. 1–22.
- Samiian, Vida (1983). *Structure of Phrasal Categories in Persian: An X-bar Analysis*. PhD thesis. Los Angeles: UCLA.
- Shokouhi, Hussein and Kipka, Peter (2003). "A Discourse Study of Persian *rā*." *Lingua* 113, pp. 953–966.
- Windfuhr, Gernot (1979). *Persian Grammar: History and State of its Study*. New York: Mouton.
- Windfuhr, Gernot (1989). "Persian." In: Comrie, Bernard (ed.), *The World's Major Languages*. London: Routledge, pp. 523–546.
- Windfuhr, Gernot (1992). "Cases." In: Yarshater, Ehsan (ed.), *Encyclopaedia Iranica* 5. New York: Bibliotheca Persica Press, 25–37.