Acta Linguistica Asiatica

TABLE OF CONTENTS

FOREWORD

Valovi	**Waves**
Nespeče morje,	Sleepless the waters,
valovi zrejo vate.	their waves laying eyes on you.
Šum zbudi srce.	Sounds inflame your heart.

          Marcel Loboda   Marcel Loboda

                                          (translation Nina Golob)

In the days of summer season openings, we are pleased to announce a new ALA issue, the second this year. Warm congratulation goes to all the authors, and feelings of gratitude go to the editorial board and all the involved. We sincerely hope that the seven research articles inflame your critical thinking just like the sounds of sleepless summer waves may inflame your hearts.

Editors and Editorial Board wish the regular and new readers of the ALA journal a pleasant read full of inspiration.

                                        Editorial board

RESEARCH ARTICLES

ASPECTS OF ASPECT: PHASIC AND EPISODIC DIMENSIONS OF VERBS

Rahman VEISI HASAR
Department of English Language and Linguistics, University of Kurdistan, Sanandaj, Iran
veisirahman@yahoo.com

Mohammad AHMADNEJAD
Department of English Language and Linguistics, University of Kurdistan, Sanandaj, Iran
m.ahmadnejad@uok.ac.ir

Abstract

The present study sets out to analyze aspectuality and coercion in Persian from a new perspective. With regard to the transcendental aspectual distinction between *perfectivity,* characterized by boundedness and heterogeneity, and *imperfectivity,* specified by uniformity and homogeneity (Langacker, 2008), it is argued that the heterogeneity of verbs may be assessed according to their phasic and episodic variables. In other words, in contrast to homogeneous verbs, which lack any kind of boundedness, heterogeneous verbs may occur either in a bounded phasic domain or in a bounded episodic domain. Concerning phasic-episodic features, this study presents a new model of lexical aspect that can differentiate five aspectual categories. The paper also scrutinizes the combinations of different verbs with different sentential operators in order to explain various kinds of type-shifting triggered by different operators. Thereby, two procedures of phasic coercion and episodic coercion are introduced which are responsible for modifying the phasic and episodic features of verbs in order to resolve the semantic conflicts between verbs and sentential operators. These procedures modify the phasic/episodic attributes of verbs according to the viewing frames evoked by interpretative operators.

Keywords: aspect; phasic and episodic coercions; adverb; tense; progressive operator

Povzetek

Študija analizira vidskost in prisilo v perzijščini iz nove perspektive. Glede na transcendentalnim vidskim razlikovanjem med dovršnostjo, za katero sta značilna leksikalna omejenost in raznoterost, in nedovršnostjo, ki jo določata enotnost in enovitost (Langacker, 2008), se domneva, da je raznovrstnost glagolov moč oceniti glede na njihove fazne in epizodične parametre. Z drugimi besedami, v nasprotju z glagoli enovitih dejanj, ki nimajo nikakršnih omejitev, se raznovrstni glagoli pojavljajo samo v omejeni domeni, ki je lahko fazna ali epizodična. V zvezi s fazno-epizodičnimi značilnostmi je v tej študiji predstavljen nov model

leksičnega vidika, ki loči pet vidskih kategorij. V prispevku so podrobneje preučene kombinacije različnih glagolov in operaterjev (udeležencev), ki razjasnijo načine spreminjanja različnih glagolskih tipov. Pri tem študija uvede dva postopka fazne prisile in epizodno prisilo, ki so odgovorni za spreminjanje faznih in epizodnih značilnosti glagolov, in na ta način razreši pomenska nasprotja med glagoli in stavčnimi operaterji. Ti postopki spreminjajo fazne / epizodne značilnosti glagolov glede na okvirje, ki jih pogojujejo interpretativni operaterji.

Ključne besede: glagolski vid; fazne in epizodične prisile; prislov; glagolski čas; progresivni operater

1   Introduction

Aspectuality lies at the intersection of ontology and linguistics owing to the fact that ontological events are represented by different verbal classes. Ontological events are classified according to attributes such as (±durative, ±telic, and ±dynamic) in the literature (e.g. Vendler, 1967; Comrie, 1976; Smith, 1997; Frawley, 1999; Moens & Steedman, 1998). These attributes demonstrate the distinctive features of events. The first attribute refers to the continuity or instantaneity of an event, the second one indicates if a process includes an endpoint or not, and the third shows whether or not the action requires a source of energy. Accordingly, five fundamental schemata are proposed which represent different ontological processes, as shown below:

1. Activity (+durative, +dynamic, -telic): walk, run, dance
2. Accomplishment (+durative, +dynamic, +telic): build a bridge, learn English, destroy a wall
3. Achievement (-durative, +dynamic, +telic): win, lose, reach to a point
4. Semelfactive (-durative, +dynamic, -telic): cough, burb, blink
5. State (+durative, -dynamic, -telic): know, love, hate

These event schemata are represented by either verbs or verbal constellations (verb phrases) in language. Although verbs partly represent these ontological categories, their aspectual natures are greatly under the influence of semantic and syntactic combinations. For example, the specificity (+/-) of the internal argument of a verb has a significant impact on the telicity feature of the verb. These two examples from Persian show this very clearly:

(1) man yek Nāme rā[1]   mi-nevisam.

  I   one Letter S/A imp-write-prs-1sg

  'I (am) write (ing) a letter.'

---

[1] *rā* in Persian indicates the specificity and the accusative case of the internal argument of a verb simultaneously. Its functions are abbreviated as S/A in this paper.

(2) Man name mi-nevisam.

    I   letter imp-write-prs-1sg

    'I (am) write (ing) letter(s).'

In the first sentence above, since the verb is linked with a specific internal argument, the verb is considered as telic (accomplishment). However, this verb in the second sentence is regarded as atelic (activity) because it has a non-specific internal argument. In other words, in the first sentence the specificity of the internal argument implies an endpoint; nevertheless, in the second sentence the action of letter writing refers to a repetitive process without implying a certain endpoint. With regard to this, some recent studies (Verkuyl, 1972, 1993, 2005; De Swart & Verkuyl, 1999), instead of focusing on event schemata and ontological categories, focus on the syntactic combinations in order to explain the aspectual nature of verbs and sentences. It must be noted that although these studies consider aspectuality as a linguistic phenomenon (not ontological), they are based on some implicit ontological categories like process, event, and state.

Nevertheless, aspectuality in cognitive grammar is based on the conceptualizer's interpretation of events rather than ontological categories and/or syntactic combinations, as is the case with the theories mentioned above. Accordingly, linguistic and ontological categories do not play a significant role in explaining the nature of aspectuality in cognitive grammar. To clarify this point, we need to further illustrate the concept of *construal*. The term construal refers to our "manifest ability to conceive and portray the same situation" in different manners (Langacker, 2008, p. 43, p. 54). Therefore, one may interpret the same scene or content in different ways and from various perspectives. Therefore, processes are not inherently characterized by aspect as being perfective or imperfective. Instead, they may be construed in different ways and with different aspects. This specific procedure of construal, which underlies aspectuality, is based on a cognitive ability that enables the construe of processual scenes in terms of heterogeneity and change (Langacker, 1987, p. 258; 2008, p. 147).

Consequently, processes may be interpreted as perfective or imperfective. An event is perfective when it is construed as a heterogeneous process revealing a change (such as a result, natural endpoint or stopping point) at the final instant of its temporal manifestation. Therefore, the process is considered as being bounded in time (e.g., build, kill, and win) simply because it is characterized by a temporal limitation. On the other hand, a process is considered as imperfective when it is construed as a homogeneous state without indicating any heterogeneity and variation (a result or a stopping point) during its temporal manifestation. Consequently, it is conceptualized as temporally unbounded, (e.g., know, hate, and like). In fact, perfective and imperfective aspects are the symptomatic manifestations of a fundamental cognitive ability specified for construing the events in terms of heterogeneity and change.

Although most events tend to be permanently categorized under one of the two perfective or imperfective aspects, it does not mean that it is impossible to interpret the same process in alternative ways; rather, the conceptualizer may sometimes interpret the same event in different ways. This is possible due to some operators called interpretative operators in this paper. These operators impose specific viewing frames on the processes denoted by verbs. As a result, they modify the aspectual nature of verbs. For instance, "-ing", as an interpretative operator in English, can re-interpret perfective verbs as imperfective. This is because this operator imposes a zooming-in strategy (internal viewing frame) on the process; accordingly, it puts emphasis on the medial phases of the event and "excludes the endpoint" and the temporal boundary of that event (Langacker, 2008, p. 155). Hence, it turns a bounded perfective verb into an unbounded imperfective one:

(3) You built the house.
(4) You were building the house.

Therefore, Langacker (2008, p. 155) believes that this operator imposes an internal perspective on perfective verbs to change them into imperfective states by removing their heterogeneity and bounded-ness. Accordingly, coercion and aspect shifting in cognitive grammar are nothing more than re-interpreting the aspect of a verbal constellation in the light of a viewing frame which has been evoked by a specific interpretative operator.

The present study aims to shed light on the relationship between the idea of heterogeneity proposed by Langacker (1982; 2008) and the five aspectual categories widely used in Vendlerian approach. In other words, we try to clarify in what ways the five aspectual categories can be inferred from the transcendental idea of heterogeneity. Next, we try to explain the role of interpretative operators in modifying the aspectual meanings of verbs in sentences. To accomplish its objectives, the paper analyzes a large amount of data in Persian language. About the data in the present study, it should be noted that the sentences have mainly been extracted from daily conversations, and in some cases, they have been constructed by the authors based on their own linguistic intuitions. Naturalness of the data and their potential interpretations have been discussed with native speakers. In addition, it is worth noting that although this theoretical study is based on data from Persian language, it seeks to propose a universal model to explain the idea of aspectuality; however, it does not claim to be universal since the validity of the findings should be assessed using data from various languages.

The article is organized as follows: In the second section, based on Langacker's model, a new model is proposed which can differentiate the five aspectual categories according to two binary variables (phasic and episodic heterogeneities). Next, in the third section of the study we try to investigate the semantic interaction between

perfective and imperfective verbs and different interpretative operators. In this regard, the concepts of phasic coercion and episodic coercion are proposed to explain this semantic interaction. These concepts refer to re-interpreting the phasic and episodic attributes of verbs in the light of viewing frames evoked by different operators. In the fourth part of the study, the semantic relationship between tense and aspect will be investigated. Finally, the fifth section of the study presents the findings of the research.

2   A cognitive appraisal of aspectual categories

As mentioned earlier, aspect is the product of a cognitive ability that conceptualizes events either as heterogeneous (perfective) or as homogeneous (imperfective) (Langacker, 1987, p. 258; Langacker, 2008, p. 147). Verbs may evoke one of these two alternative interpretations. In this regard, perfective verbs designate a heterogeneous process, which shows a kind of change and variation during its unfolding. On the other hand, imperfective verbs refer to a homogeneous and stable process not constrained by any temporal limitation. In fact, most of the verbs that are categorized under the title of imperfective aspect in cognitive grammar correspond with the ontological category of state in Vendlerian approach (example (5)). However, the perfective category covers various verbs which have been given different names (e.g., activity (example (6)), accomplishment (example (7)), achievement (example (8)), and semelfactive (example (9)) in the classical approach. In fact, cognitive grammar places these four categories of events under the perfective aspect.

(5) He knows mathematics.
(6) He walked in the park.
(7) He built a house.
(8) He won the game.
(9) He blinked.

In cognitive grammar, all divergent verbs classified under the perfective aspect have only one common attribute: heterogeneity. This feature is enough to distinguish them from the imperfective group of verbs. It should be noted, however, that most of the perfective verbs have different syntactic behaviors. In other words, they have different combinatory potential; as a result, each of them tends to combine with some particular adverbs, tenses, and interpretative operators. That is why, in Vendlerian approach, these verbs are categorized under different event schemata. For instance, accomplishment verbs preferably collocate with completive adverbs (example (10)). However, activity and semelfactive verbs are accompanied by durative adverbs (examples (11) and (12)):

(10) I cooked the meal in 2 hours.

(11) They walked in the park for 2 hours.

(12) He blinked repeatedly for 2 hours.

In spite of the subtle differences between different perfective verbs, they all have in common the attribute of heterogeneity; all of them profile events as heterogeneous processes. However, the concept of heterogeneity is too broad to explain the subtle differences of these verbs. The main question is how it is possible to extract the four aspectual categories (activity, accomplishment, achievement, and semelfactive) from this fundamental concept. In order to answer this question, we need to explain the concept of heterogeneity in such a way that it can explain the common attribute to all the four different perfective verbs, and explicate their characteristic features. In the following lines, we try to discuss this precise and comprehensive conception of heterogeneity in detail.

Considering Langacker's model (2008, pp. 147-149), one may argue that perfective verbs are characterized by phasic and/or episodic heterogeneities. As far as phasic heterogeneity is concerned, it refers to the emergence of a result (a change) at the final instant of atemporal period in which an event is unfolding. The result, opposed to the the previous co-equal phases of the process, gives rise to a kind of heterogeneity. This specific sort of heterogeneity is called *phasic heterogeneity* since "some change" (Langacker, 2008, p. 147) in the phasic structure[2] of the process "is observed". In other words, the preliminary particles of a process give birth to an ultimate result (as the final change) in an evolutionary manner.

In fact, this attribute can distinguish between perfective verbs (achievement and accomplishment) characterized by the presence of phasic heterogeneity (result) and those specified by the absence of phasic heterogeneity (without result) (activity and semelfactive). Thus we can say that perfective verbs can be divided into two groups of resultative (+R) and result-less (-R) based on the phasic heterogeneity variable (phasic variable). Smith (1977) describes the result emerging from the phasic structure of a process as *natural* ending. The word "natural" is used to differentiate accomplishment and achievement verbs from activity and semelfactive ones which lack an output. The following sentences in Persian illustrate some verbs including phasic heterogeneity that end in a result.

(13) Man xāne   Rā   sāxtam.

    I   house S/A build-pst-1sg

    'I built the house.'

---

[2] Phasic structure refers to the structure of an event which includes preliminary, medial, and final phases.

(14) ʔu   bāzi   Rā   bord.
     s/he game   S/A win-pst-3sg
     'She won the game.'

In the examples above, there is exactly a single output for each resultative process. In other words, each process as a whole gives birth to one result. The singularity of the result is the fundamental requirement for the phasic heterogeneity since it is only by highlighting a single result (as the most prominent element) against the other co-equal components of the process (other phases of the process) that the conceptualizer may perceive a phasic heterogeneity. Therefore, we say that events entail phasic heterogeneity only if they have a single result against the preliminary phases. This is called the singularity requirement of the result. However, if the result is repeated, the (+R) feature will disappear. That is to say, when there are multiple outputs for a single process, that process is considered as result-less. This is because there is not a singular resultant stage to be in sharp contrast with other phases.

(15) ʔu   bāzi   Rā   bord.
     s/he game S/A win-pst-3sg
     'S/he won the game.'

(16) ʔu   bāzi   Rā   čandin bār   bord.
     s/he game S/A several times win-pst-3sg
     'S/he won the game several times.'

In the first example above (15), the event of winning ends in a result that is the same for all stages and elements of the process. This distinguished result is a prominent heterogeneous element that is in contrast to the previous phases (stages) of the event. On the other hand, the process in the second example (16) is a repetitive process (repetitive activity). As a result, there are several results for numerous realizations of winning, and consequently, several results can be expected for the whole process. The verb in this example is thus considered to be (-R), and is not different from a perfective semelfactive verb.

(17) ʔu   čandin bār   čeʃmak zad.
     s/he several time Blink   lv.pst-3sg
     'S/h blinked several times.'

In fact, an event is (+R) and has phasic heterogeneity only if there is one and only one result which is in contrast to the previous phases of the event. The singularity requirement and its relation to the phasic heterogeneity can be stated as follows:

One state can be the result of a process if and only if this state is the only result for all the stages (phases) and for all the components of the process, and if it is *different* from the other elements.

In fact, an element or a state can be considered as the *result,* creating phasic heterogeneity, if it is the only output for all the phases of the event.

Nevertheless, Langacker (2008, p. 148) believes that there is a different kind of heterogeneity that does not relate to the resultant stage. He states that although some processes do not include a result and, consequently, do not have any phasic heterogeneity, yet they are categorized under perfective aspect. He argues that these processes can be regarded as heterogeneous with regard to their *episodic* dimensions. In fact, the episodic aspects of a process may give rise to a specific kind of heterogeneity which is called episodic heterogeneity. To explain this, firstly, we need to clarify the episodic dimensions of the perfective verbs, and then investigate how these dimensions bring about episodic heterogeneity. As far as episodic dimensions are concerned, they refer to the internal and external boundaries of a perfective event. To explain this, let us compare countable-uncountable dichotomy with perfective-imperfective duality. In the literature, this comparison has always been enlightening and very helpful to grasp the concept of aspect (see Leech, 1970, 2004; Mourelatos, 1978; Hoepelman & Rohrer, 1980; Carlson, 1981) and may clarify the notion of boundary and its relationship to episodic heterogeneity. With regard to the nominal distinction, countable nouns (e.g., table) refer to entities that are bounded and have a discrete boundary with their surroundings. In other words, an *external boundary* separates the entity from the spatial environment. This is while uncountable nouns (e.g., water) indicate elements that do not have a clear boundary with their surroundings since they are amorphous and shapeless. On the other hand, a countable noun refers to an entity with an internal configuration. Constituents of such entities are related to each other in a specific arrangement, and further, constituting particles can be differentiated from each other. Therefore, *internal boundaries* (lines of demarcation and separation) between the entity's particles (that are gathered in a structured arrangement) may be assumed. However, an uncountable noun denotes a shapeless element that does not have a discrete configuration, and its internal particles do not form a stable structure. That is why its constituents cannot be distinguished from each other. In other words, there are no clear and stable internal boundaries and lines of separation between internal elements of this entity. In short, countable nouns refer to elements that have a clear boundary line with their surroundings (external boundary), and there exists a clear and stable line of distance between their internal particles (internal boundary) that are gathered together in a structured configuration. However, neither the external limitation nor the internal boundary are relevant to an uncountable noun due to their external and internal shapelessness.

In comparison with the nominal distinction, perfective and imperfective verbs can be distinguished from each other according to the ideas of internal and external boundaries. Similar to uncountable nouns, the imperfective verbs profile processes that are devoid of any kind of external boundary with their temporal surroundings[3]. Accordingly, they "exclude" endpoint and temporal limitation "from what" they "put onstage for focused viewing" (Langacker, 2008, p. 147). In other words, these processes do not convey an ending or limitation; therefore, they are not bounded in time. Similar to uncountable nouns, imperfective processes do not contain recognizable internal phases or medial particles. In other words, a static process is devoid of medial phases and middle particles (the absence of internal boundary) since it is free of any kind of structured internal configuration. As a result, a state remains the same for all the instants during its temporal manifestation (Vendler, 1967, pp. 106-107; Smith, 1997, p. 32; Langacker, 1982, p. 266; Binnick, 1991, p. 183). Thus imperfective processes, similar to uncountable elements, have neither external (endpoint) nor internal boundaries (lines of demarcation between the internal phases of the event). However, most perfective events, similar to countable entities, have both external boundary (bounded-ness) and internal limitation (internal configuration). The external boundary separates the event from the temporal surrounding. This outer limitation is described either as a natural endpoint or as a stopping point. Due to this temporal limitation, perfective events are described as temporally bounded (Langacker, 2008, p. 147). In addition, the perfective processes involve discernible internal phases and different medial particles. Therefore, there are lines of demarcation between the components and stages of a perfective process.

To further elaborate on this issue, first, let us discuss the external boundaries of perfective events. Contrary to static processes, which have removed from *focused viewing* all boundaries with their temporal surroundings, all perfective events directly or implicitly express an endpoint (the completion or a stopping point). These events are thus described as temporally bounded. Temporal boundedness is evident in a perfective event with phasic heterogeneity because the moment in which a result emerges is the ending point of that action and the indicative of the boundary between the action and the temporal surrounding. In other words, the ultimate result of a perfective process may separate the event from the temporal surrounding and turn it into a bounded process. The outer limitation made by the result is called external phasic boundary. As already mentioned, it is this external phasic limitation that evokes phasic heterogeneity and boundedness. This suggests that the result as the final stage in a perfective process illustrate an external boundary.

Apart from the external boundary made by the emergence of a result, Langacker (2008, p. 148) suggests the idea of episodic bounded-ness (events "*occurring in*

---

[3] By temporal surrounding, we mean the infinite temporal axis in which a bounded event emerges. In fact, the bounded event only occupies a limited part of this infinite line.

bounded episode") to explain the external limitation of a result-less perfective process (e.g., walk, swim, blink and burp). In this regard, Langacker makes a comparison between some specific countable nouns (lake, lawn, brick, beep, and hole) and result-less perfective verbs. Despite the fact that these entities (like a lake) are inherently homogeneous, their external boundaries with the surroundings (a line between the lake and the land) can be recognized. In other words, they are episodically bounded. In the same way, although some events do not involve the resultant stage as the external limitation, they have an episodic external boundary separating the event from the temporal surrounding. Smith (1997, p. 23) believes that although a result-less process does not complete and finish at the resultant stage, it stops at a specific point described as *stopping point*. This point illustrates a specific kind of external limitation separating the bounded event without result from its temporal surrounding. While the result emerging necessarily from the phasic structure of a resultative process, the stopping point is derived from the terminable nature of perfective events with no results.

In contrast to the protractible nature of imperfective processes that removes any kind of external limitation, the terminable nature of result-less events imposes a limitation on these processes. The terminability attribute implies a specific temporal point in which a terminable perfective process may *come to* stop. Consequently, the result-less events are bounded due to this point. Therefore, if we regard the completion point of a resultative process as external phasic boundary, it can be argued that the stopping point of a result-less event may be considered as external episodic limitation. With regard to the terminable attribute of result-less processes and their external episodic limitations, the result-less events tend to always be bounded. That is why verbs *like walk, swim, blink and burp,* contrary to imperfective verbs, easily combine with the progressive operator (-ing) and *resist the present tense* (Langacker, 2008, pp. 147-148). In other words, since they associate the property of terminability, they always tend to be bounded. Similarly, result-less perfective verbs can always be combined with the progressive operator in Persian (example (18)); however, the imperfective ones cannot collocate normally with this operator (example (19)):

(18) Sohrāb dāʃt   qadam mi-azd.
      Sohrab pro-pst-3sg walk   imp-lv-pst-3sg
      'Sohrab was walking.'

(19) *kāmbiz dāʃt   fāteme rā   dust dāʃt.
      Kambiz   pro-pst-3sg fateme S/A love lv.pst-3g
      ('Kambiz was loving Fateme.')

We call this boundary, which demonstrates the event's bounded-ness and its detachment from its temporal surrounding, external episodic boundary. This episodic

boundary brings about episodic bounded-ness and heterogeneity. The heterogeneity or change made by a stopping point or an episodic limitation is described as episodic heterogeneity. This episodic heterogeneity is made by an external limitation and is thus called external episodic heterogeneity. Nevertheless, the heterogeneity made by the result which functions as the external phasic boundary is called external phasic heterogeneity. All kinds of perfective verbs are always bounded because there is a demarcation line (episodic or phasic external boundary) that separates these events from temporal surroundings. In other words, all the bounded perfective processes include an external limitation. This outer boundary as the common attribute in all perfective verbs may emerge in different forms such as a result or a stopping point.

A conceptualizer may demonstrate its cognitive access to the external boundary of an event in different ways. For instance, using the perfect operator (past participle (verb + e) + ʔastan (to be)), the conceptualizer may highlight the external phasic limitation (result) or the external episodic boundary (stopping point) of perfective verbs. The function of this interpretative operator is to focalize the external limitations of perfective verbs. A combination of this operator with various perfective verbs looks quite normal since all of them have external boundaries either in a phasic form or in an episodic form.

(20) Man zyād   david-e ʔam.
     I   very much run-pp. be-prs-1sgs
     'I have run very much.'

(21) ʔu   xāne rā   sāxt-e   ʔast.
     s/he house S/A build-pp. be-prs-3sg
     'S/he has built the house.'

(22) ʃomā   bāzi   rā   bord-e   ʔi.
     you-pl game S/A win-pp. be-prs-2pl
     'You have won the game.'

(23) Man ʔatse   kard-e   ʔam.
     I   sneeze lv-pp.   be-prs-1sg
     'I have sneezed.'

In the examples above, external boundaries have been highlighted by the perfect operator. In fact, the operator has emphasized the external phasic limitation of the resultative perfective verbs (examples (21) and (22)), and the external episodic boundary of the result-less verbs (examples (20) and (23)). However, when this operator combines with a static and homogeneous process, usually the resulting

construction is a marked combination (example (24)) because the states do not have external boundary to be highlighted by the operator. That is because, as mentioned earlier, imperfective verbs delete the external episodic boundary from viewing; therefore, they are not used with perfect construction.

(24) *to ryāzi   balad bud-e ʔi.
      you mathematics know be-pp. be-prs-2sg
      ('You have known mathematics.')

Nevertheless, it should be noted that some static verbs combine with this operator just in virtue of episodic coercion. This means that they obtain new episodic features under the influence of a perfect operator, or , an external episodic limitation is imposed on the static process by the perfect operator (example (25)).

(25) ʔu   tā   ʔalān mariz bud-e ʔast.
      s/he until now sick   be-pp. be-prs-3sg
      'S/he has been sick until now.'

Phasic and episodic coercion phenomena will be discussed in detail in the next section.

There are some specific verbs (described as super-lexical morphemes by Smith (1997, p. 48)) profiling either the external phasic boundary or the outer episodic limitation. For instance, using some verbs in Persian, *motavaqef kardan* 'stop' and *tamām kardan* 'finish', a conceptualizer may profile the external limitation of a perfective process as the focus of attention. The first operator is used to highlight the external episodic boundary (stopping point) and the second one is applied to focalize the external phasic one (completion point):

(26) To qadam zadan rā   motavaqef kardi.
      you walking lv   S/A stop   lv-pst-2sg
      'You stopped walking.'

(27) Man xāne   sāxtan rā   tamām kardam.
      I   house building S/A finish   lv-pst-1sg
      'I finished house building.'

Nevertheless, static verbs that are devoid of any phasic and episodic external boundaries do not combine with these super-lexical morphemes. In other words, they do not have external boundaries to be highlighted and emphasized by these verbs:

(28) *Hasan moʔalem budan rā   motavaqef/tamām kard.
     Hasan   teacher   being S/A stop/finish   lv-pst-3sg
     ('Hasan stopped/finished being teacher.')

(29) *Maryam mariz budan ra   motavaqef/tamām kard.
     Maryam sick   being S/A stop/finish   lv-pst-3sg
     ('Maryam stopped/finished being sick.')

However, this combination may convey an interpretation only if the verb undergoes phasic coercion. In this case, the verb describes a change or transformation from one state to another. This specific construction in which the state verbs undergo phasic coercion will be addressed in the next section.

It can be claimed that all perfective events have either external episodic boundary or external phasic boundary, and such limitations can be highlighted by different operators. These external boundaries bring about a specific kind of external heterogeneity that refers to a change (a result or a stopping point) obtained at the final temporal point of the temporal extension. The heterogeneity made by a result is called phasic heterogeneity, and that which is made by episodic limitation (stopping point) is called episodic heterogeneity. However, these different external boundaries (phasic and episodic) have the same function as far as they impose an outer limitation on the processes and bring about bounded-ness and perfectivity.

Now, let us discuss the internal boundaries of perfective events. Accomplishment and activity verbs, similar to countable nouns, refer to processual entities having discernible and distinguishable internal stages and particles. That is, the medial boundaries between the constituting elements of the process are distinguishable. Therefore, by using the progressive operator (*dāʃtan*), the conceptualizer can highlight some of the internal particles of a process against the external boundary (the endpoint or the stopping point of the action).

(30) Karim dāʃt   xāne mi-sāxt.
     Karim pro-pst-3sg house imp-build-pst-3g
     'Karim was building a house.'

(31) ʔebi dāʃt   qadam mizad.
     Ebi   pro-pst-3sg walk   imp-lv-pst-1sg
     'Ebi was walking.'

In fact, the role of the progressive interpretative operator is to highlight the medial phases and elements of a process (Langacker, 2008, p. 155) and exclude the endpoints.

Since these two processes (accomplishment and activity) have accessible and discernable internal particles, their combination with the progressive operator looks completely normal. Consequently, in the examples above, the medial stages and particles of building and walking processes have been focalized, and the endpoints (the result or the stopping point) have been de-emphasized and excluded by the operator. That is because the progressive operator has highlighted some of the medial particles against the external boundary of the actions. Nevertheless, because static verbs designate processes without discernible and accessible medial stages and internal elements, they cannot combine with the progressive operator of *dāʃtan*:

(32) *to   dāʃti   ryāzi   balad budi.
     You pro-pst-2sg mathematics know   be-pst-2sg
     ('You were knowing mathematics.')

(33) *man dāʃtam   Ɂu   rā   dust dāʃtam.
     I   pro-pst-1sg s/he S/A love lv-pst-sg
     ('I was loving her/him.')

In fact, separability and recognizability of the internal elements of an event indicate the internal boundary (or existence of demarcation lines) between the event's internal phases and particles. These demarcation lines cause a kind of episodic heterogeneity between the constituents of the event. The fact that the medial particles (elements) of an event are cognitively accessible to the conceptualizer is called *internal boundary accessibility*. It means that the episodic boundaries between the medial elements of the event are accessible. Also, the resulting episodic heterogeneity made by this is called *internal episodic heterogeneity*. In other words, the episodic lines of demarcation between different elements of a process may give rise to an episodic heterogeneity. In fact, this specific episodic heterogeneity only exists in events which have accessible and recognizable medial elements.

In the following lines we discuss the internal configuration and boundaries of achievement and semelfactive verbs. As discussed earlier, these two events contain external phasic boundary and outer episodic limitation respectively, and this leads to the emergence of a kind of external heterogeneity in them. However, it should be noted that the internal particles and medial phases of these events do not look accessible to the interpretative operators. Thus, when the verbs that represent these events are combined with the progressive operator, the operator cannot highlight the internal stages and particles of these events, but it highlights the other parts of the events.

(34) ʔu   dāʃt   bāzi rā   mi-bord.
     s/he pro-pst-3sg game S/A imp-win-pst-3sg
     'S/he was winning the game.'

(35) Man dāʃtam   ʔatse   mi-kardam.
     I   pro-pst-1sg sneeze imp-lv-pst-1sg
     'I was sneezing.'

For instance, in the first example above, the interpretative operator has highlighted the internal elements of the preparatory phase, which has been imposed as a new episodic feature[4] on the process by the progressive operator (*dāʃtan*) (example (34)). The same operator, combined with a semelfactive event, has highlighted the external boundaries between several realizations of sneezing (example (35)). In other words, since the internal stages and particles of this event are inaccessible, the operator focalizes the external limitations between consecutive events [5] with inaccessible internal components. In conclusion, the medial particles have been conceptualized in neither case. In previous approaches (Vendler, 1967; Binnick, 1991; Smith, 1997; Frawley, 1999) these two categories (achievement and semelfactive) have been illustrated as punctual or instantaneous to show that they do not contain internal particles. The present study does not concern whether or not these events, from an ontological or physics' viewpoint, contain internal boundaries or particles. We only, from a cognitive perspective, claim that even if these events possess internal particles, they are not accessible to the conceptualizer, and thereby they do not have internal boundary accessibility. It means that their internal particles and the boundaries between them do not undergo cognitive processing. Therefore, it can be argued that although these two processes have external heterogeneity because of the existence of episodic/phasic external boundaries, they are devoid of internal episodic heterogeneity since their medial elements are inaccessible.

The examples below clearly show the difference between events with accessible internal boundaries and events with inaccessible internal boundaries.

---

[4] The verb undergoes episodic coercion in this example (see example (61)).

[5] Due to the semantic conflict between the episodic feature of the semelfactive event and the viewing frame evoked by the progressive operator, the semelfactive verb exceptionally expresses a repetitive process including several realizations of sneezing (episodic coercion). In this regard, the operator puts emphasis on the external limitations between the successive events (numerous realizations of sneezing) characterized by inaccessible internal particles (see example (62)).

(36) Do sāʔat maŝghul-e   neveŝtane ʔān name budam   vali hanuz
     For 2 hour engaged-Ez   writing   that letter be-pst-1sg but yet

     tamām   naŝode.
     finish   not-become-prs-3sg

     'I was engaged in writing that letter for two hours, but it has not been finished yet.'

(37) Man be-moddate do sāʔat ʔatse   kardam.
     I   for   two hour sneeze lv-pst-1sg
     'I sneezed for two hours.'

In the first sentence above (example (36)), which contains an accomplishment verb with accessible medial particles, the middle elements of the process correspond with the moments of the temporal extension described by the durative adverb. However, because a semelfactive event does not have any accessible internal particles, there is not such a correspondence. The second example on the other hand shows the correspondence of temporal moments indicated by the adverbial operator with the external episodic boundaries between numerous realizations of sneezing. This sentence clearly points to a repetitive activity including several instantiations and realizations of sneezing, and each external boundary between two sub-instantiations corresponds with a moment of the temporal extension described by the adverbial operator. The reason is that semelfactive events do not have accessible internal elements, and their external episodic boundaries are only accessible to the operator. As a result, external boundaries between semelfactive events are mapped onto the temporal moments of the adverbial temporal extension.

Accordingly, we say that semelfactive and achievement events only possess accessible external boundaries and are devoid of internal boundary accessibility. Since static events are devoid of internal boundaries and accessible elements, there happens no correspondence between different moments of the temporal extension (expressed by the adverb) and the internal elements of the static process. The state stays the same for all the instants because its internal structure is inaccessible and unanalyzable for temporal moments:

(38) ʃahlā   tā se ruz mariz bud.
     Shahla for 3 day sick   be-pst-3sg
     'Shahla was sick for three days.'

Although the adverbial operator has imposed an external episodic limitation on the verb by episodic coercion (see examples (83)-(86)), the state as a whole stays the same for all the moments of the temporal extension. Therefore, the differential moments do

not correspond with the medial phases (that are inaccessible) of the static process. Based on this introduction, the following rule for cognitive access to internal and external episodic boundaries can be presented.

An event is episodically accessible if the conceptualizer is able to make a contrast between a part of the event as the primary focus and the remaining stages. The highlighted element can be a medial element (accessible internal boundaries) or the external boundary (accessible external boundary: a result or a stopping point).

Briefly said, episodic heterogeneity in perfective events is the product of their external and internal boundaries. If the external boundary (a result or a stopping point) of an event is accessible, the event will have external heterogeneity and external boundary accessibility. Moreover, if the internal boundaries of an event are accessible, the event will have internal heterogeneity and internal boundary accessibility.

Based on the phasic (-R, +R) and episodic features [internal boundary accessibility (α), external boundary accessibility (β)], we can clearly categorize perfective events. The binary phasic variable examines events' phasic heterogeneity, and the binary episodic variable examines events' episodic heterogeneity. The phasic variable takes two complementary features (-R and +R), but the episodic variable is a variable with two non-complementary features (α and β). The episodic features are non-complementary because the co-existence of these values in a same process is not impossible. The same process cannot be simultaneously resultative (+R) and resultless (-R), but it may contain both of internal (α) and external (β) boundary accessibilities. Before classifying different perfective verbs according to the features mentioned above, it should be noted that since all perfective verbs are constrained by an external limitation (a result or a stopping point), the feature β cannot differentiate the perfective verbs from each other. In other words, they all have the feature + β in common because it is afundamental feature which makes a distinction between perfective verbs and static ones. The external limitation of the perfective processes appears in the forms of a result or a stopping point. As a result, for classifying the perfective verbs, only the feature of result (R) and the attribute of internal boundary accessibility (α) are used as distinctive features.

As far as the phasic variable is concerned, the processes can be divided into +result and −result. This variable can differentiate the accomplishment and achievement verbs from activity and semelfactive ones. In addition, with regard to internal episodic feature (α), the processes are classified into +α (events with accessible internal boundaries) and −α (events with inaccessible medial particles). This parameter makes a distinction between accomplishment and activity verbs and semelfactive and achievement ones. In other words, only activity and accomplishment verbs contain internal boundary accessibility (+α), but semelfactive and achievement verbs are characterized by inaccessibility of internal boundaries (-α).

To sum up, different perfective verbs have a common feature (+ β), but they are different with regard to the phasic attribute (+/-R) and internal episodic feature (+/-α). It should be noted though that all of these different features (phasic and episodic ones) are nothing more than different manifestations of the fundamental idea of heterogeneity. The table below describes the aspectual features of perfective events according to the fundamental attributes of phasic and episodic features. It shows how the four fundamental categories of lexical aspect can be extracted from the cognitive idea of heterogeneity. These features not only analyze perfective events in a unified manner but also differentiate them from each other.

Table 1: Aspectual categories according to phasic and episodic variables

Aspectual categories	Phasic heterogeneity variable	Episodic heterogeneity variable
Accomplishment	+R (resultative)	+ α (internal boundary accessibility) + β (external boundary accessibility)
Activity	-R (resultless)	+ α (internal boundary accessibility) + β (external boundary accessibility)
Achievement	+R (resultative)	- α (lack of internal boundary accessibility) + β (external boundary accessibility)
Semelfactive	-R (resultless)	-α (lack of internal boundary accessibility) + β (external boundary accessibility)

According to the table above, we can evaluate the degree of phasic and episodic heterogeneities of different perfective verbs. It can be inferred that accomplishment events have the greatest degree of heterogeneity. However, semelfactive events have the least amount of heterogeneity because they only contain external episodic heterogeneity. It should be noted that the feature + β is the minimum requirement for a verb to be considered as perfective. The other two events contain moderate degrees of heterogeneity:

Continuum of heterogeneity:
Accomplishment > Activity, Achievement > Semelfactive

The above mentioned features suggest that a static process is a process with no phasic or episodic bounded-ness, and that it has neither internal episodic boundary nor external phasic/episodic limitations:

State: -α, -β, -R

In fact, this process removes the external phasic and episodic limitations from the focused viewing. In addition, its medial phases (internal episodic limitations) are inaccessible. The only difference between a state and a semelfactive verb is due to the presence of external episodic limitation in the latter. Based on this introduction, in the

next part we discuss the semantic interaction between the discussed aspectual categories and different interpretative operators at the sentential level.

3   Aspectual categories and interpretative operators in Persian

As already mentioned, interpretative operators impose specific viewing frames on the event schemata denoted by verbs. For instance, perfect and progressive operators put emphasis on the external phasic-episodic boundary and the internal episodic limitations of events respectively, or shortly, they re-construe the processes indicated by verbs in the light of specific viewing frames. Temporal adverbs may also play the role of interpretative operators in the sentence; because the aspectual features of verbs are re-construed according to the temporal scales evoked by adverbs (Langacker, 2008). Consequently, verbs are highly sensitive to the temporal features of adverbs, which is because aspectual features of verbs are re-conceptualized in terms of the qualitative (evolutionary or non-evolutionary temporal extension) and quantitative (durative or non-durative) features of adverbial temporal scales. Combining a verb and a certain adverb, the conceptualizer makes a relationship between the temporal scale of the adverb and the phasic-episodic attributes of the verb.

Durative and completive adverbs are among those adverbs that exert considerable influence on the aspectual meanings of verbs. Before delving into this semantic relationship, we should illustrate the interpretative functions of these adverbs in detail. As far as a durative adverb is concerned, it re-interprets the phasic-episodic features of a verb in the light of a homogeneous temporal extension without culmination and result. However, the completive adverb re-construes the phasic and episodic features of a verb in terms of a heterogeneous temporal extension, which is directed towards a result. With regard to a durative adverb, the temporal moments hold the same degree of importance in the homogeneous temporal extension, and the process is supposed hide any kind of quality change over this consistent and uniform temporal extension. However, the completive adverb demonstrates a heterogeneous temporal extension with qualitatively different moments appearing in an evolutionary frame. The temporal extension gives rise to a result (ultimate change) in an evolutionary manner.

Since adverbs convey some qualitative and quantitative attributes, they are highly selective when it comes to compatible verbs. They prefer to be combined with those verbs that have some well-matched features. In case of incompatibility between the temporal nature of an adverbial scale and the event schema indicated by a verb, coercion will come forth to remove the incongruity. Coercion, in the literature, is a procedure responsible for resolving semantic incompatibility between different linguistic elements in a sentence (see Pustejovsky, 1995; Moens & Steedman, 1998, p. 17; Smith, 1997, p. 48; De Swart, 1998, p. 349; 2002; De Swart & Verkuyl, 1999; Jackendoff, 1997; Michaelis, 2004, 2011; Cortés-Rodriguez, 2014). As mentioned

before, coercion in cognitive grammar refers to re-construing the aspectual features of verbs in the light of the viewing frames invoked by operators. As a result, some phasic or episodic features of a verb may change due to the semantic requirements of an operator in the sentence. As far as the semantic interaction of verbs and adverbs is concerned, the episodic features (α/β) and the phasic attributes (+R/-R) of the verbs may undergo modification under the influence of the adverbial operators.

On the other hand, different aspectual features (phasic and episodic dimensions) may be modified in the process of coercion. In this regard, coercion may modify the phasic attributes of verbs; and alternatively it may change the episodic dimensions of verbs. As a result, the procedure of coercion triggered by different operators can be divided into two categories: phasic and episodic coercions. For instance, the phasic coercion, which is triggered by an adverb, modifies the phasic attribute of a verb according to the temporal features of the adverbial scale, and the episodic coercion, motivated by an adverb, may change the episodic properties of a verb according to the temporal requirements of the adverbial operator.

We next try to investigate the semantic relationship between the aspectual categories and different adverbs (completive and durative) in different sentences. The following sentences are mainly in the past tense because the past tense is neutral towards any kind of aspect, and triggers no coercion in Persian. By choosing the past tense, we can remove the disturbing variable of tense in order to examine the semantic interaction of verbs with adverbs.

We first study the relationship between perfective verbs and different operators, and then investigate the relationship between imperfective verbs and different operators. The following examples illustrate various constructions made by various perfective verbs and different adverbs:

(39) ʔali yek borj   rā   sāxt.
    Ali one tower S/A build-pst-3sg
    'Ali built a tower.'

(40) ʔali dar do sāʔat yek borj   rā   sāxt.
    Ali in 2 hour one tower S/A build-pst-3sg
    'Ali built a tower in 2 hours.'

(41) *ʔali be moddate do sāʔat ʔan borj   rā   sāxt.
    Ali   for   2 hours that tower S/A build-pst-3sg
    ('Ali built that tower for 2 hours.')

(42) ʔali be moddate do sāʔat borj   sāxt.

Ali for   2   hours tower build-pst-3sg

'Ali built tower for 2 hours.'

(43) ʔānha qadam zadand.

They walk   lv-pst-3pl

'They walked.'

(44) ʔānhā be moddate do   sāʔat qadam zadand.

They for   two hour walk   lv-pst-3pl

'They walked for 2 hours.'

(45) *ʔali dar do sāʔat qadam zad.

Ali   in 2 hour walk   lv-pst-3sg

('Ali walked in 2 hours.')

In example (39) the accomplishment verb (*sāxtan*), with a specific internal argument, indicates an event which is characterized by + R and + (α/β). The process shows a certain result, and its internal and external boundaries are accessible. The combination of this verb with a completive adverb in example (40) sounds natural. It is because the phasic attribute of the accomplishment verb (+R) is completely compatible with the heterogeneous and evolutionary temporal extension indicated by the completive adverb. On the other hand, the episodic property of the verb (+ α: accessible internal boundaries) is in harmony with the durative temporal scale of the adverb. The accessible internal boundaries make it possible to make a connection between internal components of the event and different instants of the temporal extension of the adverb. The internal elements can thus correspond to different instants of the temporal extension, and after the combination of this verb and the adverb, the phasic and episodic properties of the verb remain intact. However, the combination of this verb and the durative adverb in example (41) is semantically marked. Although the episodic property of the verb (+α) is compatible with the durative temporal extension of the adverb, its phasic attribute (+R) is incompatible with the homogeneous and non-evolutionary temporal extension of the adverb. As such, the sentence is marked. Nevertheless, this verb has a non-specific internal argument in example (42) and therefore the process expressed by the verb is not directed towards a certain result or a singular output but rather to a repetitive activity of tower building. It is considered as an activity verb (the activity of building tower as such) and is characterized by aspectual features of –R and + (α/β). Since the verb refers to a result-less process, it can normally be combined with a durative adverb which indicates a homogeneous and non-

evolutionary temporal scale. The verb with the phasic attribute of (-R) is therefore in harmony with the homogeneous temporal extension of its adverb.

In example (43), the activity verb expresses a process specified by properties -R and + (α/β). A combination of this verb and the durative adverb is natural in example (44) because the phasic feature of the activity verb (-R) is congruent with the homogeneous temporal extension of the adverb. On the other hand, the episodic attribute of the verb (+ α) is harmonious with the durative temporal extension of the adverb. As a result of the accessibility of internal boundaries, internal components of the process indicated by the verb correspond with the moments of temporal extension profiled by the adverb. This means that the verb has kept all of its aspectual properties without any modification. However, the combination of this verb is incompatible with the completive adverb in example (45) because the phasic attribute of the activity verb (-R) is in conflict with the evolutionary and heterogeneous temporal extension of the adverb; therefore, this sentence is marked and unnatural. However, this verb in combination with the completive adverb can make sense if it describes a situation in which a person has a plan for walking in a specific day for three hours but, one day and because of some reasons, he has to finish this plan in less than two hours. In this special context and under the influence of the completive adverb, the process is re-construed as a heterogeneous process with a natural endpoint (end of the plan) which is obtained at the final temporal point of the evolutionary temporal extension denoted by the adverb. Accordingly, the activity verb seems to be re-interpreted as an accomplishment verb in order to be compatible with the evolutionary temporal scale expressed by the adverb. In other words, the –R feature of the verb is turned into + R by the adverb. It can be argued that the activity verb in this example undergoes phasic coercion because its phasic attribute changes into + R. This kind of coercion is considered as a kind of phasic coercion since the phasic attribute of the verb (-R) has been modified (+R) according to the temporal requirements of the adverbial operator (evolutionary time).

When these adverbs are combined with achievement and semelfactive verbs, many complicated constructions are obtained. As mentioned earlier, semelfactive verbs are characterized by aspectual properties of -R and -α/+β, and achievement verbs are characterized by +R and -α/+β. The following examples illustrate the combinations of these verbs and different adverbs:

(46) Sara bāzi   rā   bord.
     Sara game S/A win-pst-3sg
     'Sara won the game.'

(47) Sara dar   do   sāʔat bazi rā   bord.
     Sara in   2   hour   game S/A   win-pst-3sg
     'Sara won the game in 2 hours.'

(48) Sara be moddate do   sāʔat bāzi rā   bord.
     Sara for   2   hour   game S/A   win-pst-3sg
     'Sara won the game for 2 hours.'

(49) To   cheŝmak zadi.
     You blink   lv.pst.2sg
     'You blinked.'

(50) To   be moddate do sāʔat cheŝmak zadi.
     You for   2   hour   blink   lv-pst-2sg
     'You blinked for 2 hours.'

(51) *ʔu dar do sāʔat cheŝmak zad.
     s/he in   2   hours blink   lv-pst-3sg
     ('S/he blinked in 2 hours.')

Example (46) includes an achievement verb with the features of +R and -α/+β. However, this verb is combined with a completive adverb in example (47). This combination which describes a situation in which Sara has won the game after 2 hours of struggle. In this example, the heterogeneous and result-oriented temporal extension indicated by the adverb is in harmony with the phasic attribute of the achievement verb (+R). However, durative dimension of the completive adverb is in contrast with the inaccessibility of the internal episodic boundary of the verb (-α), which means that the internal components of the verb can not be accessed by the durative temporal scale of the adverb, and as a consequence the internal elements may not correspond with the moments of the temporal extension. To remove this semantic conflict, a durative preparatory phase with accessible internal particles (a new episodic feature: + α) is added to the process, which makes different instants of the adverbial temporal extension correspond with the accessible internal components of the preparatory phase of the process of winning. The preparatory phase is extended to two hours because it is the only relevant part of the process that is accessible to the completive adverb. As such, the process characterized by inaccessibility of internal particles changes into a durative process (two hours of duration in the preparatory phase) with accessible internal elements, and the ultimate result will be obtained after this temporal period. It can be argued that the achievement verb has taken a new episodic attribute (+ α), and although the internal elements of the process were inaccessible at first, they become accessible under the influence of the adverbial operator. As a result, the verb seems to have undergone episodic coercion because its episodic aspect (-α) changes into +α, however the phasic attribute stays the same during the process of coercion.

In example (48) the achievement verb combines with a durative adverb. In this case, the homogenous and durative temporal extension of the adverb is completely incompatible with the phasic (+R) and episodic aspects (-α) of the verb. This sentence is meaningful only in a specific context in which the person has repeatedly won a certain game for two hours. The verb, in this sense, describes a repetitive activity; simply because the process lacks a singular result for all of its components (it has several results). It can be argued that the achievement verb changes into a repetitive activity to be compatible with the semantic requirements of the durative adverb. This verbal metamorphosis can resolve the phasic and episodic conflicts between the verb and the adverb as the repetitive dimension of the resulting verb simultaneously removes the phasic (+R of the verb) and episodic (-α of the verb) incompatibilities. If we consider the repetitive activity as an eventual complex, the sub-events (several realizations of winning) will be its internal components. In addition, the internal components of the eventual complex are accessible to be connected with the instants of the durative temporal extension expressed by the durative adverb. At this place we can argue that the achievement verb, with inaccessible internal elements (-α), is turned into a repetitive activity with accessible internal components (+α). This aspectual modification illustrates an episodic coercion which is triggered by the durative adverb. Because the repetitive activity lacks a singular result for all of its elements and phases, it cannot be considered as a resultative verb. Consequently, the verb with + R turns into a –R verb in the light of the adverb with a non-evolutionary temporal extension. This aspectual modification demonstrates a phasic coercion motivated by the adverb, and procedures take place to modify the eventual schema denoted by the verb according to the temporal properties of the adverb.

Example (49) includes a semelfactive verb with aspectual features of –R and -α/+β. This verb is combined with a durative adverb in example (50). Although the homogenous temporal extension of the adverb is compatible with the phasic attribute of the verb (- R), the durative aspect of the temporal extension profiled by the adverb is inharmonious with the inaccessibility of the internal components of the process (-α, +β). As a result of this semantic conflict, the semelfactive verb turns into a repetitive activity. The resulting repetitive activity includes multiple co-equal instantiations (numerous realizations of blinking). If we consider the repetitive activity as a process, the sub-events (different realizations of blinking) can be regarded as its accessible internal components, and the realizations of the repetitive activity would correspond with the instants of the durative temporal extension of the adverb. With regard to this point, it can be said that the semelfactive verb takes a new episodic feature (+α: accessible internal particles) by the episodic coercion just to be harmonious with the temporal scale of the adverb; however, the phasic attribute remains unchanged in this process. The episodic coercion re-construes the episodic features of the verb in the light of temporal attributes of the adverb operator.

Example (51) includes a semelfactive verb which is combined with a completive adverb. In this case, the durative and heterogeneous dimensions of the adverb are in contrast to the phasic (- result) and episodic (-α, +β) attributes of the verb. The internal components of the process expressed by the verb are inaccessible to the temporal extension of the adverb, and the result-less verb can not be combined with an adverb indicating an evolutionary temporal frame. However, this sentence is acceptable in a situation in which the speaker had been trying to blink for two hours, and finally he succeeded in blinking just once after two hours. In this case, the verb takes new episodic and phasic attributes to be compatible with the temporal requirements of the completive adverb. For instance, the verb takes a new preparatory phase that is supposed to result in an output. In this regard, the imposed preparatory phase has two functions: first, for adding a durative phase with accessible internal components to the event that is to be processed against the durative temporal period indicated by the adverb, and second, to add a result to the result-less process which is to be re-construed against the evolutionary temporal extension of the completive adverb. A specific durative phase has been added to the process for resolving the episodic conflict between the verb and adverb. In other words, the verb has acquired a new episodic attribute (+α: preparatory phase with accessible elements) to be compatible with the durative temporal extension of the adverb. In addition, it takes a new phasic property (+R) to be compatible with the heterogeneous temporal scale of the adverb, and the verb simultaneously undergoes phasic and episodic coercions to modify its aspectual features according to the temporal requirements of the adverb.

The conclusion that arises is that all kinds of type shifting are imposed because of a conflict between the phasic-episodic attributes of the verbs and the temporal configurations of the adverbs. In case of a phasic or episodic conflict, two coercion procedures are applied to resolve this conflict. The phasic coercion is applied to re-construe the phasic attribute of the verb according to the temporal features of the adverb. Similarly, the episodic coercion is imposed to modify the episodic aspects of the verb according to the temporal properties of the adverb.

This evokes our further investigation on the relationship between two operators (namely perfect and progressive operators) and different verbs in different sentences. In the following examples, the interaction between the perfect operator and different verbs are illustrated:

(52) Man ʔān tasivir rā keŝid-e   ʔam   / budam.
     I   that picture S/A draw-pp. be-prs-1sg / draw-pp. be-pst-1sg
     'I have drawn that image.' / 'I had drawn that image.'

(53) Man qadam zad-e ʔam   / budam.
     I   walk   lv-pp. be-prs-1sg / be-pst-1sg
     'I have walked.' / 'I had walked.'

(54) ʔu   be xāne   resid-e   ʔast   / bud.
     s/he to home   reach-pp.   be-prs-3sg / be-pst-3sg
     'S/he has reached the home.' / 'He had reached the home.'

(55) Man cheŝmak zad-e ʔam   / budam.
     I   blink   lv-pp. be-prs-1sg. / be-pst-1sg.
     'I have blinked.' / 'I had blinked.'

(56) Man farad   qabl ʔaz ʔāmadane to   dars   rā   tamām kard-e ʔam.
     I   tomorrow before coming-Ez you lesson S/A end   lv-pp.
     'I will have finished the homework before you come.'

The important point in the above examples is that the combination of different aspectual categories with the perfect operator sounds normal. In examples (52)-(56), the perfective verb highlights the external episodic or phasic boundary of the processes. Since all perfective verbs have external boundaries, they can be combined with the perfect operator. In resultative processes (accomplishment and achievement), the perfect operator profiles the result (the external phasic boundary) as the focus of attention. However, in result-less processes (semelfactive and activity) the perfect operator profiles the external episodic boundary (β) as the focus of attention. The perfect operator functions as a cognitive adjusting operator providing the conceptualizer with access to some specific parts (the ultimate result or the external episodic boundary) of the events. Because all perfective processes have either a final result (+R) or an episodic boundary (β), they are in harmony with the viewing frame evoked by this operator. No coercion occurs in these examples.

In the following lines, the interaction of the imperfective operators (progressive and habitual operators) and different aspectual classes will be investigated. It can be argued that the imperfective operator is the opposite of the perfect one. In fact, while the perfect operator highlights the external limitation of the processes, the imperfectivizers defocus the external episodic and phasic boundaries of the events. More specifically, the imperfective operators change the perfective processes into unbounded states. Frawley (1999, p. 328) describes the imperfective as an open aspectual category due to the lack of phasic and episodic limitations in imperfectivized processes. In Persian there are two different kinds of imperfectivizers (*dāŝtan and mi-*

): the first one (*dāŝtan*) homogenizes [6] the processes by representing them as progressive actions, but the second one (*mi-*[7]) homogenizes the processes by turning them into habits. It can be argued that these interpretative operators impose different viewing frames on the event schemata denoted by verbs. As far as *dāŝtan* as a progressive operator is concerned, it imposes an internal viewing frame (zooming-in strategy) on the processes; it highlights the medial components of events and defocuses their external phasic-episodic limitations. Nevertheless, *mi-* as a habitual operator imposes an infinite viewing frame (zooming-out strategy) on processes, and erases their internal and external phasic-episodic limitations:

(57) Man dāŝtam   qadam mi-zadam.
     I   pro-pst-1sg walk   imp-lv-1sg
     'I was walking.'

(58) ʔu   hamiŝe qadam mi-zanad.
     s/he always walk   imp-lv-3sg
     'S/he always walks.'

Similar to -ing in English, *dāŝtan* as a progressive operator imposes an internal perspective on processes by selecting an "internal portion" of the process "for focused viewing" in order to "exclude the endpoints" (Langacker, 2008, p. 155). This operator shows the image of being in the middle of a process by highlighting the internal elements and removing the external boundaries of the processes. Although the external limitation of a perfective operator is defocused under the influence of this operator, its internal episodic limitations remain unchanged. It is because the medial elements and their episodic limitations are highlighted by the progressive operator. For instance, some medial elements and middle phases of the walking process in example (57) are focalized by the operator; however, the external episodic limitation of this action is defocused.

Mi- operator nevertheless imposes an infinite viewing frame on the processes, and consequently removes the external phasic and episodic boundaries of events and turns them into unbounded states. The habit of walking in example (58), for example, does not convey a stopping point but describes a state without any external limitation. This external unboundedness is not the side effect of focalizing the medial components as is the case with the progressive operator; on the contrary, it is the result of re-construing the perfective verb in the light of zooming-out strategy evoked by the

---

[6] Homogenization is a process in which the heterogeneity of a verb is deleted by an operator.

[7] mi- is a polysemic prefix. It can be used as progressive and habitual operators. Being a progressive operator, it behaves completely like dāŝtan. Accordingly, we ignore this semantic aspect of *mi* in this paper.

habitual operator. The habitual operator does not focalize medial elements of the process but removes their internal limitations, which consequently makes the resulting static process stay the same for all the instants (Smith, 1997, p. 32). In other words, since the perfective events are turned into static processes by the habitual operator, they do not have internal configuration (internal boundaries) to be connected with the instants of the temporal extensions. Considering example (58), walking doesn't describe an action with accessible particles, ongoing in some specific temporal moments, on the contrary it describes a state without internal configuration which remains the same for all the instants. Finally, it should be noted that *dāŝtan* operator only defocuses the external phasic and episodic boundaries while *mi-* deletes both internal and external boundaries. *Mi-* can in that way be regarded as an absolute homogenizer while *dāŝtan* is considered as a restricted homogenizer. It can be argued that *mi-* as a habitual operator changes the whole aspectual features of processes by changing events into states that lack any kind of limitation; nevertheless, *dāŝtan* as a progressive operator modifies only the external limitations of events by defocusing them. The following examples illustrate the combination of dāŝtan operator with different verbs:

(59) To   dāŝtai   xāne   ra   mi-saxti.
     You   pro-pst-2sg house S/A   imp-build-pst-2sg
     'You were building the house.'

(60) Ɂānhā dāŝtand   qadam mi-zadand.
     They   pro-pst-3pl walk   imp-lv-pst-3pl
     'They were walking.'

(61) Ɂu   dāŝt   bāzi   rā   mi-bord.
     He   pro-pst-3sg   game   S/A   imp-win-pst-3sg
     'He was winning the game.'

(62) Man dāŝtam   Ɂatse   mi-kardam.
     I   pro-pst-1sg sneeze imp-lv-pst-1sg
     'I was sneezing.'

In examples (59) and (60), sentences include accomplishment and activity verbs characterized by aspectual features + R and + (α/β) and –R and + (α/β) respectively. They are combined with *dāŝtan* imperfectivizer in a normal way. The middle components of these processes are cognitively accessible to be highlighted against the external phasic and episodic boundaries. External phasic-episodic boundaries are defocused in favor of a highly focused partial process. In other words, the operator

highlights some specific medial components of the processes, and deemphasizes their external episodic and phasic limitations. Characterized by defocused external boundaries, the processes are homogenized to a considerable degree. Besides, the above mentioned verbs do not say anything about their possible endings (the stopping point or the result) because their external limitations have been removed by the operator. Actually, activity and accomplishment verbs undergo episodic and phasic coercions because their external episodic and phasic limitations are deemphasized by the operator. As far as the accomplishment verb in example (59) is concerned, it undergoes episodic and phasic coercions because its external episodic and phasic boundaries are removed by the operator; however, the activity verb in example (60) undergoes episodic coercion since its external episodic limitation is defocused under the influence of the operator. These verbs undergo these coercions to be compatible with the viewing frame invoked by the operator. It is worth mentioning that the internal episodic limitations of the above mentioned processes remain intact during the coercion procedures triggered by the progressive operator. Therefore, the medial particles are focalized by the operator to profile the imagery of *being-between-a-process*.

As already mentioned, the primary function of the progressive operator is to focalize medial elements. It can be said that this operator can not highlight middle components of the processes with inaccessible internal particles ($-\alpha$). In example (61), there is an achievement verb characterized by $+ R$ and $-\alpha/+\beta$. Medial parts of this process are not cognitively accessible. Consequently, there is an episodic conflict between the verb and operator. For resolving this conflict, a durative preparatory phase (with accessible particles) is added to the achievement process. As a result, the operator highlights some specific middle parts of the preparatory phase and defocuses the external phasic-episodic boundary of the process in order to represent the process as a progressive one. It can further be said that the verb has acquired a new episodic feature ($+\alpha$: preparatory phase with accessible components) to be compatible with the internal viewing frame evoked by the operator. In addition, external phasic boundary of the process is removed by the operator and thus the verb does not indicate anything about the possible ending of the process. It can be argued that the verb undergoes episodic coercion (preparatory phase addition) and phasic coercion (defocused external phasic limitation) to be congruent with the viewing frame evoked by the operator. These coercions are applied as the result of re-construing the achievement verb in the light of the internal perspective evoked by the operator.

In example (62), there is a semelfactive verb including features $-R$ and $-\alpha/+\beta$. As in the previous case, the middle components of the process are not accessible for the operator. There is a semantic inharmony between the episodic feature of the verb and the cognitive function of the operator which is used for profiling the internal components, which makes the semelfactive verb turn into a repetitive activity to be harmonious with the progressive operator. If we regard the repetitive activity as an

eventual complex, the sub-events (several realizations of sneezing) will be its internal components. In such case the progressive operator can highlight some internal particles (some medial realizations) of this repetitive activity. The semelfactive verb characterized by – α is episodically coerced into a repetitive activity specified by + α in order to be harmonious with the perspective evoked by the progressive operator. In addition, the external episodic limitation (the last realization of the action) of the repetitive activity is defocused by the progressive operator. It is because the operator puts focus only on the medial particles of the repetitive activity. Therefore, it can be said that this verb undergoes two episodic coercions to be compatible with the internal perspective evoked by the operator.

Dāŝtan as a progressive operator modifies the phasic and episodic features of processes to bring about a homogenized process. To accomplish its objective, this operator defocuses the external phasic and episodic limitations of the processes, and focalizes some medial episodic limitations. This operator erases the external phasic and episodic heterogeneity by deleting the outer phasic and episodic limitations. Nevertheless, the internal episodic heterogeneity remains intact since the internal boundaries are highlighted by the operator.

However, the second imperfectivizer (prefix of *mi-*) evokes a zooming-out strategy; it is applied to re-construe a process as a habitual state. The operator excludes the internal and external phasic-episodic boundaries of a process to represent it as an open ended state. The following examples illustrate the interaction of this operator with various verbs.

(63) Man in Ɂāwāz rā hamiŝe sobhā   mi-xānam.
     I   this song S/A always morning-PL imp-sing-prs-1sg
     'I sing this song every morning.'

(64) Ɂu hamiŝe pyāderavi mi-konad.
     He always walk   imp-lv-prs-3sg
     'He always walks.'

(65) Time mā hamiŝe bāzi   rā   mi-bāzad.
     Team our always game S/A imp-lose-3sg
     'Our team always loses the game(s).'

(66) Ɂu   hamiŝe Ɂatse   mi-konad.
     s/he always sneeze imp-lv-prs-3sg
     'S/he always sneezes.'

In the above examples, there are different predicates denoting various processes with different phasic and episodic boundaries. However, *mi-* operator has homogenized all of these perfective verbs. For example, the verb in example (63) has the features +R and + (α/β). Nevertheless, when it is combined with a *mi-* operator, its phasic and episodic boundaries are not accessible any longer. The verb in this sentence refers to the persistent continuum of a habit without external phasic-episodic limitation. It seems that the habitual operator removes the external phasic limitation of the process to represent it as an infinite state (phasic coercion), and that the imperfectivized process as a state does not include internal configuration to be connected with different moments of a temporal extension. Accordingly, the state remains the same for all the instants. Here it can be argued that the internal components of the process are not accessible any longer (episodic coercion). Then, the accomplishment verb loses its phasic and episodic features to be compatible with the infinite viewing frame evoked by the habitual operator. This absolute homogeneity is provided by phasic and episodic coercions triggered by the habitual operator. Similarly, the verb in example (64) involves the features -R and + (α/β). But, it is turned into a habit by the *mi-* operator. Like in the previous case, its middle components and its external episodic limitation are not accessible anymore. In other words, the verb undergoes episodic coercions to be compatible with the infinite viewing frame evoked by the operator, which makes the internal and external episodic limitations of the process get removed in order to be compatible with the zooming-out strategy of the operator. In the same way, the verbs in examples (65) and (66) (achievement +R and -α/+β; semelfactive -R and -α/+β) do not refer to specific processes with certain episodic and phasic boundaries. Instead, they describe persistent states without being bound to a phasic or an episodic limitation. In such cases the external episodic boundary of the semelfactive event is removed by the operator (episodic coercion) to illustrate the event as an infinite state. Similarly, the outer phasic limitation of the achievement process is deleted by the operator (phasic coercion) to represent the process as a persistent habit. Consequently, this operator turns all different aspectual categories into homogeneous states by applying phasic and episodic coercions. In other words, it deletes the internal/external phasic and episodic boundaries of the perfective verbs.

As mentioned earlier, interpretative operators mainly modify the phasic and episodic features of perfective verbs. But, how do they combine with imperfective verbs that have no phasic-episodic heterogeneity? For instance, the perfect operator highlights the external phasic and episodic limitations of the perfective verbs. Static verbs, characterized by having no external limitation, on the other hand do not seem to be compatible arguments for the perfect operator. As a result, many of the stative verbs in the perfect construction are semantically marked, especially when they describe an inherent attribute and a persistent state:

(67) *ʔu ryāzi   balad bude ʔast.
     s/he mathematics know be-pp. be-prs-3sg
     ('She has known mathematics.')

(68) *ʔu ʔāŝeqe setāreh bude   ʔast.
     s/he love   Setareh be-pp. be-prs-3sg
     ('S/he has loved Setareh.')

However, episodic coercion may resolve this combinatory incongruency, particularly, when the static verbs describe a transient or non-inherent state (see Croft, 2012, p. 56). When the perfect operator combines with these stative verbs, the resulting construction illustrates a transient state which is constrained by an external episodic boundary. Accordingly, rather than being considered as infinite states, they are regarded as transient states preceded or followed by an episodic limitation. As a result, example (68) may be considered as a meaningful sentence only if it describes a transient psychological state which is not necessarily true at the present moment. The following examples demonstrate this construction.

(69) ʔu   mariz bude ʔast.
     s/he sick   be   prs-per-3sg.
     'S/he has been sick.'

(70) ʔu   dar ʔāmricā bude ʔast.
     s/he in America be   prs-per-3sg.
     'S/he has been in America.'

(71) ʔānhā ʔasabāni bude ʔand.
     They angry   be   prs-per-3pl
     'They have been angry.'

To remove the semantic conflict between the perfect operator and stative verbs, the perfect operator turns the states into processes characterized by an external episodic boundary. This means that stative verbs are conditioned by an outer episodic limitation, from which it can be argued that stative verbs may acquire an external episodic limitation in virtue of the episodic coercion triggered by the perfect operator. Although these verbs have a kind of external episodic boundary, they can not be combined with progressive operators:

(72) *Ɂānhā Ɂasabāni mi-bude-Ɂand.

They   angry   imp-be-pp.-be-prs-3pl

('They had been being angry.')

(73) *Ɂānhā dārand Ɂasabāni bude-Ɂand.

They   imp-3pl angry   be-pp.-prs-3pl

('They have been being angry.')

Although the states are conditioned by external episodic boundary in perfect construction, they do not include accessible internal components to be highlighted by the progressive operator. This construction describes a transient state which remains the same for all the instants of a limited temporal extension. In other words, it illustrates a transient state without internal phases and particles which holds for a restricted temporal period.

Similarly, imperfective verbs do not combine with the progressive operator in a normal way. The following examples illustrate this problematic combination in the past and present tenses:

(74) *man dāram/dāŝtam ryāzi   balad hast-am/bud-am.

I   pro-prs/pst-1sg mathematics know lv-prs/pst-1sg

('I am/was knowing mathematics.')

(75) *Ɂānhā dārand/dāŝtand be tārix   Ɂalāqemand hastand-budand.

They   pro-prs/pst-3pl to history like   lv-prs/pst-3pl

('They are/were liking history.')

(76) *to dāri/dāŝti   mariz hasti-budi.

You pro-prs/pst-2sg sick   lv-prs/pst-2sg

('You are/were being sick.')

As the above examples show, the stative verbs can not be combined with *dāŝtan* imperfectivizer simply because they do not have internal components to be highlighted against the external limitation. However, some stative verbs can be connected with *dāŝtan* operator only by virtue of coercion. That is to say, the imperfective verb takes a perfective interpretation to remove the semantic incompatibility between the stative verb and the restricted imperfectivizer.

(77) ʔu   ʔaz   sib   badeŝ   mi-ʔumad.
     S/he from apple bad-3sg imp-lv-pst-3sg
     'S/he disliked apple.'

(78) ʔu   dāŝt   ʔaz   sib   badeŝ   mi-ʔumad.
     S/he pro-pst-3sg from apple bad-3sg imp-lv-pst-3sg
     'S/he was disliking (gradually) apple.'

Although the verb in example (77) is a stative verb describing a psychological inherent state, it refers to the process of taste changing in example (78). In other words, it describes a psychological process in which one kind of personal taste is in the process of changing. Accordingly, the stative verb is turned into a progressive one containing internal particles (durative preparatory phase) some of which are highlighted by the progressive operator. Therefore, it can be argued that the verb takes a new episodic feature ((+α): accessible internal boundaries) to be harmonious with the zooming-in strategy evoked by the operator. Accordingly, we argue that this verb acquires internal heterogeneity (accessible internal episodic limitations) by undergoing episodic coercion triggered by the operator. Although the verb in example (78) describes a process with accessible internal components, it does not say anything about its possible outputs. Because progressive operator excludes any kind of external phasic-episodic limitation from conceptualization, it does not add a new phasic limitation to the verb. Therefore, an active phasic limitation is not present in the structure of the resulting process.

By the same token, stative verbs do not combine with completive adverbs, because the homogeneous imperfective states are incompatible with the heterogeneous temporal extension of completive adverb:

(79) *ʔu   dar do sāʔat ryāzi   balad ʔast/bud.
     S/he in   2   hours mathematics know be-prs/pst-3sg
     ('S/he knows/knew mathematics in two hours.')

(80) *ʔu   dar 2 ruz   mariz ʔast/bud[8].
     S/he in   2 days sick   be-prs/pst-3sg.
     ('S/he is/was sick in two days.')

---

[8] dar is not considered as a locating adverb in this sentence.

However, the completive adverb may change a stative verb into a resultative process.

(81) ʔunā 2 sāʔate ʔin   ro   dunestan.
     They in 2 hours this fact S/A know-pst-3pl.
     'They knew this fact in two hours.'

In example (81), the completive adverb transforms the stative verb into a resultative verb specified by two hours of preparatory phase (episodic coercion) and an acquired result (phasic coercion). In other words, the state process undergoes phasic and episodic coercions to be compatible with the evolutionary temporal extension described by the adverb. Similarly, some adverbs (like: *kam kam* 'gradually', *dir* 'late' and *yedafʔe* 'suddenly') may change some stative verbs into resultative ones by phasic coercion:

(82) Man yekdafʔe/kam kam dunestam   ke   ʔun xāʔene.
     I   suddenly/gradually know-pst-1sg that s/he traitor-be-3sg
     'I suddenly/gradually knew that s/he is a traitor.'

Finally, let us consider the combination of stative verbs and the durative adverb. Although stative verbs, indicating an inherent attribute, do not tend to combine with durative adverbs (example (83)), the transient ones can easily go with them (examples (84) and (85)).

(83) *ʔu do sāl (e) ryāzi   balad bud/ʔast.
     s/he 2   year (be) mathematics know be-pst/prs-3sg
     ('S/he knew/knows mathematics for two years.')

(84) ʔu   do sal   mariz bud/ʔast.
     s/he 2   year sick   be-pst/prs-3sg
     'S/he was/is sick for two years.'

(85) To   do sāʔat (e)   ʔasabani budi/hasti.
     You 2   year (be) angry   be-pst/prs-2sg
     'You were/are angry for two hours.'

As far as the durative adverb is concerned, it indicates a homogeneous temporal extension which begins at one point and ends at another. Whenever it combines with a state, it imposes external episodic limitations on the infinite state. As a result, the static process acquires a beginning point (left external episodic limitation) coinciding

with the first moment of the adverbial temporal extension; additionally, it takes a stopping point (right outer episodic boundary) coinciding with the final instant of the adverbial temporal period. The resulting construction describes a state which holds for a limited temporal period. Therefore, the verb undergoes episodic coercions to be harmonious with the temporal requirements of the adverbial operator. Although this static verb is characterized by external episodic limitations, its internal components are not accessible. Consequently, the resulting construction does not go with *dāŝtan* as a progressive operator:

(86) *ʔu dāŝt   do sal   mariz bud.
    s/he pro-pst-3sg (for) 2   year sick   be-pst-3sg
    ('S/he were being sick for two years.')

Then, it can be argued that the static process in this construction remains the same for all the instants of the temporal extension. In other words, it does not have internal components to be connected with the moments of the temporal period.

Accordingly, it can be said that although the states exclude any kind of internal and external episodic-phasic limitations from conceptualization, they may acquire phasic and episodic boundaries to be compatible with the viewing frames evoked by different interpretative operators. In this regard, phasic and episodic coercions are applied to modify the aspectual features of verbs according to the semantic features of the operators. The phasic coercion may add phasic heterogeneity to the states, while the episodic coercion can add episodic heterogeneity to the static processes.

Finally, it is worth mentioning that light verbs play a significant role in indicating aspectual meanings in Persian. The light verb of *budan* 'be' (*ʔast* in the present and *bud* in the past) indicates an imperfective meaning (example (87)). Also, *ŝodan* 'become' as a light verb is used to refer to phasically heterogeneous perfective processes (example (88)). Additionally, the light verb of *kardan* 'make' is mainly used for episodically heterogeneous perfective verbs (example (89)).

(87) To   ʔŝeq budi.
    You lover lv-pst-2sg
    'You were in love.'

(88) To   ʔāŝeq ŝodi.
    You lover   lv-pst-2sg
    'You fell in love.'

(89) ʔu   ʔāŝeqi   kard.

S/he love-ness lv-pst-2sg

'S/he /He did loving.'

Accordingly, when there is semantic incompatibility between verbs and interpretative operators, the procedure of light verb changing can also be used to resolve the semantic incongruency. In the next section, the relationship between tense and aspectual features will be scrutinized in detail.

4   Tense and Aspect

In contrast to aspect which illustrates the internal temporal constituency of a situation, tense as a deictic phenomenon makes a relationship between the temporal location of a situation and the deictic moment of speaking (Comrie, 1985). The deictic moment of speech is applied to determine and locate the temporal positions of the events. From a cognitive perspective, this deictic center functions as the reference point for the situation as the target (for reference point model see, Langacker 1995; 2009). In fact, it provides the conceptualizer with mental access to the temporal position of the event.

Point of reference (point of speech)   >>>   target (point of situation)

Time as represented by tense system is divided into two spheres (past and present) according to the deictic moment of speech (t0). The present sphere contains three temporal zones (pre-present, present and post-present) that are anchored around t0; however, the past sphere which includes only one temporal zone (past zone) is completely disjoined from t0. The past tense precedes the moment of speaking (t0), and, in contrast to the other tenses, it is completely disconnected from that moment (Declerck, 2006, 2015). The distance and separation of the past tense from the punctual moment of speech (t0) put the immediate temporal scope completely before t0 in the past tense. Consequently, the past temporal zone has a sufficient temporal scope to "encompass a bounded event" (Langacker, 2009, p. 191). Therefore, the past tense can represent processes in perfective and imperfective ways (Langacker, 2009, p. 190). Just in the same way, verbs in Persian can be expressed as perfective and imperfective in past tense:

(90) Mādar qadam zad.   / madar dāŝt   qadam mi-zad.

Mother walk   lv-pst-3sg / mother pro-pst-3sg walk   imp-lv-pst-3sg

'Mother walked.' / 'Mother was walking.'

(91) ʔu   ʔāwāz rā   xānd.   / ʔu   dāŝt   ʔāwāz rā   mi-xānd.
      s/he song S/A sing-pst-3sg / s/he pro-pst-3sg song   S/A imp-sing-pst-3sg
      'S/he sang the song.' / 'S/he was singing the song.'

(92) ʔānhā bāzi   rā   bordand.   / ʔānhā dāŝtand   bāzi rā   mi-bordand.
      They game S/A win-pst-3pl / they   pro-pst-3pl play S/A imp-win-past-3pl
      'They won the game.' / 'They were winning the game.'

(93) ʔu   cheŝmak zad.   / ʔu   dāŝt   cheŝmak m-izad.
      s/he blink   lv-pst-3sg. / s/he pro-pst-3sg blink   imp-lv-pst-3sg
      'S/he blinked.' / 'S/he was blinking.'

(94) ʔu   ʔali rā   dust dāŝt.
      s/he Ali S/A love have-pst-3sg
      'S/he loved Ali.'

As far as present tense is concerned, it locates "the situation as coinciding" with the punctual moment of speech (Declerck, 2006, p. 173). Bounded events should correspond to t0 to convey the meaning of being in the present zone, however, concerning default viewing arrangement and impossibility of temporal coincidence between a bounded process and t0 in present tense (Langacker, 2009, p. 192, 2011; Declerck, 2006), verbs tend *not* to convey perfective meaning in present tense. In other words, because heterogeneous bounded events can not be coincided with the punctual moment of speaking in the present tense (Declerck, 2006, p. 173), most verbs in present tense are often expressed as imperfective[9]. However, perfective verbs can be used in present tense construction if they are homogenized by the imperfectivizer operator (Declerck, 2006, p. 174). Therefore, perfective verbs are homogenized into progressive or habitual verbs in the present tense construction.

Similarly, verbs should always be imperfective in present tense in Persian language. As a result, verbs are always attached to *mi-* as an imperfective marker in the present construction. In fact, the prefix *mi* can never be separated from perfective verbs in the present tense; otherwise the structure is ungrammatical:

(95) *ʔu   hamin hālā davad.
      S/he just   now run-prs-3sg
      ('S/he is running right now.')

---

[9] There are some specific and marked constructions in which perfective verbs can be expressed in present tense. However, because they are not the primary concern of this paper we do not discuss them here.

The obligatory presence of *mi-* in the present construction illustrates symptomatically the fact that verbs should always be imperfective (progressive or habit) in the present temporal zone. Hence, *mi-* is an indispensable part of the present tense construction in Persian. This prefix as an imperfective operator homogenizes perfective verbs in the present tense. As a polysemous prefix, *mi-* can express progressive and habitual meanings.

Persian perfective verbs are homogenized by a polysemic prefix to meet the requirements imposed by the default viewing frame. According to this frame, which is evoked by the present tense, the coincidence between the speech and the bounded event is impossible (problem of duration). In addition, the conceptualizer doesn't have an epistemic dominance over a process which is in progress at the present moment. As a result, it can not view the external phasic-episodic limitations of the process (epistemic problem) (Langacker, 2009, pp. 191-192; 2011). To resolve the epistemological and episodic conflict between the *viewing frame* evoked by present tense and the *perfective verbs*, Persian language applies the prefix *mi-* to remove the external boundaries of bounded processes denoted by perfective verbs.

As far as the progressive meaning of *mi-* in the present tense construction is concerned, it deletes the external phasic and episodic boundaries of the processes, and highlights some medial phases. Therefore, internal perspective is imposed on the processes to remove the semantic incompatibility between the present tense and the perfective verbs:

(96) ʔu   hamin hālā mi-davad.

     s/he just   now imp-run-prs-3sg

     'S/he is running right now.'

(97) ʔu   hamin hālā xāne   rā   mi-sāzad.

     s/he right   now house S/A imp-build-prs-3sg

     'S/he is building the house right now.'

(98) ʔu   hamin hālā bāzi   rā   mi-barad.

     s/he right   now game S/A imp-win-prs-3sg

     'S/he is winning the game right now.'

(99) ʔu   hamin hālā cheŝmak mi-zanad.

     s/he right   now blink   imp-lv-pr-3sg

     'S/he is blinking right now.'

Verbs in the present tense can be homogenized by a restricted homogenizer (*mi-* as progressive marker) to remove the inharmony between perfective verbs and the present tense. In this case, *mi-* behaves completely like dāŝtan as a progressive operator, in such a way that it emphasizes the medial particles, and defocuses the external limitations. Therefore, it can be argued that the present tense construction may impose an internal viewing frame on processes. The verbs undergo phasic and episodic coercions to be harmonious with the temporal and epistemological perspective invoked by the present tense construction.

However, *mi-* as a habitual operator can change perfective verbs into habitual states in the present tense. In this construction, *mi-* excludes both external and internal phasic-episodic boundaries of the processes. Consequently, verbs don't describe bounded actions in the *present* moment; on the contrary, they describe some habitual states that are not constrained by t0 and any other episodic-phasic limitation. In this case, *mi-* as an absolute homogenizer erases any kind of internal/external phasic and episodic limitations of the processes. In other words, perfective verbs undergo phasic and episodic coercions triggered by *mi-* as an absolute homogenizer in order to be compatible with the present tense:

(100) Ɂu   hamiŝe sigār   mi-keŝe.
      s/he always cigarette imp-lv-prs-3sg.
      'S/he smokes.'

(101) Ɂu   hamiŝe pyāderavi mi-konad.
      s/he always walk   im-lv-prs-3sg
      'S/he always walks.'

Accordingly, the bounded verbs in the present tense are homogenized in two ways: firstly, the conceptualizer may exclude the external episodic-phasic boundaries of events by using a restricted homogenizer. Second, s/he can delete both internal and external limitations of processes by applying an absolute homogenizer. Therefore, the present tense turns verbs into imperfective verbs which lack external limitations.

5   Conclusion

In this paper, we tried to study lexical categories and the phenomenon of coercion from a new perspective. Following Langacker (2008), we divided verbs into two groups, namely perfective and imperfective. The perfective category describes a heterogeneous process which is bounded in time, but the imperfective one illustrates homogeneous and unbounded processes which are not constrained by any external

limitation. However, the heterogeneity of the perfective verbs may be assessed according to phasic and episodic variables. In other words, processes can be regarded as perfective, when they occur in a bounded phasic domain (with a single result highlighted against the preparatory and middle phases) or in a bounded episodic domain (having internal or external episodic boundaries). Afterwards, we classified perfective verbs into four types according to the phasic and episodic variables. In addition, the imperfective verb as the fifth aspectual type lacks any kind of heterogeneity.

Table 2: Aspectual categories according to phasic and episodic variables

Aspectual categories	Phasic heterogeneity variable	Episodic heterogeneity variable
Accomplishment	+R (resultative)	+ α (internal boundary accessibility) + β (external boundary accessibility)
Activity	-R (resultless)	+ α (internal boundary accessibility) + β (external boundary accessibility)
Achievement	+R (resultative)	- α (lack of internal boundary accessibility) + β (external boundary accessibility)
Semelfactive	-R (resultless)	-α (lack of internal boundary accessibility) + β (external boundary accessibility)
State	-R (resultless)	-α (lack of internal boundary accessibility) -β (lack of external boundary accessibility)

When there is a semantic conflict between the aspectual features of verbs and the viewing frames evoked by interpretative operators, phasic and episodic coercions are applied to remove the semantics incompatibility. These procedures modify the phasic and episodic attributes of verbs according to the semantic requirements of interpretative operators. These coercions refer to re-construing the event schemata denoted by verbs in the light of the viewing frames evoked by operators.

In addition, we introduced two different kinds of imperfectivizers of *mi* and *dāŝtan* in Persian. We showed that *dāŝtan* as a progressive operator highlights the internal boundaries of a perfective process and excludes the external phasic/episodic boundaries. But *mi* as a habitual operator defocuses both internal and external phasic/episodic boundaries in favor of representing the process as an infinite state. Accordingly, the first one was called a restricted homogenizer and the second one was called an absolute homogenizer. In contrast to the homogenizers, which defocus the external boundaries of events, it was argued that the perfect operator highlights the

external phasic/episodic boundaries of processes. Finally, we showed that verbs can be expressed as perfective and imperfective in the past tense because the past temporal zone is completely separated from t0. However, perfective verbs should be imperfectivized in the present tense because the bounded event cannot be coincided with the punctual moment of speaking. This homogenization can be accomplished in two ways: first, the conceptualizer may exclude the external phasic/episodic boundary of the bounded event and highlight the internal components. Second, it may defocus both internal and external phasic-episodic boundaries of the bounded process.

References

Binnick, R. (1991). *Time and the verb: A guide to tense and aspect.* Oxford: Oxford University Press.

Carlson, L. (1981). Aspect and quantification. In P. J. Tedeschi & A. Zaenen, *Syntax and semantics: Vol. 14. Tense and Aspect* (pp. 31-64). New York: Academic Press.

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.

Cortés-Rodríguez, F. J. (2014). Aspectual features in role and reference grammar: A layered proposal. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, *27*(1), 23-53.

Croft, W. (2012). *Verb: Aspect and causal structure.* New York: Oxford University Press.

Declerck, R. (2006). *The grammar of the English tense system: A comprehensive analysis*. Berlin and New York: Mouton de Gruyter.

Declerck, R. (2015). *Tense in English: Its structure and use in discourse.* London: Rutledge.

De Swart, H. (1998). Aspect shift and coercion. *Natural language & linguistic theory*, *16*(2), 347-385.

De Swart, H., & Verkuyl, H. (1999). Tense and aspect in sentence and discourse. *Lecture notes of the 11th European summer school in logic, language and information.* Utrecht Institute of Linguistics OST.

De Swart, H. (2000). Tense, aspect and coercion in a cross-linguistic perspective. In *Proceedings of the Berkeley Formal Grammar conference*. University of California, Berkeley.

Frawley, W. (1999). *Linguistic semantics*. London: Routledge.

Hoepelman, J. and Rohrer, C. (1980). On the mass count distinction and the French imparfait and passe simple. In C. Rohrer (Ed.), *Time, tense, and aspect* (629-645). Tuebingen: Niemeyer.

Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.

Langacker, R. W. (1982). Remarks on English aspect. In P. J. Hopper (Ed.), *Tense-aspect: Between semantics and pragmatics* (pp. 265-304). Philadelphia: John Benjamin Publishing.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). California: Stanford University Press.

Langacker, R. W. (1995). Possession and possessive constructions. In J. R. Taylor, & R. E. MacLaury (Eds.), *Language and the cognitive construal of the world* (pp. 51-80). Berlin: Mouton de Gruyter.

Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.

Langacker, R. W. (2009). *Investigations in cognitive grammar.* Berlin: Mouton de Gruyter.

Langacker, R. W. (2011). The English present: Temporal coincidence vs. epistemic immediacy. In A. Patard, & F. Brisard (Eds.), *Cognitive approaches to tense, aspect, and epistemic modality* (pp. 45-86). Philadelphia: John Benjamins Publishing Company.

Leech, G. (1970). *Towards a semantic description of English*. Bloomington: Indian University Press.

Leech, G. (2004). *Meaning and the English verb*. London: Longman.

Michaelis, L. A. (2004). Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive linguistics*, *15*(1), 1-68.

Michaelis, L. A. (2011). Stative by construction. *Linguistics*, *49*(6), 1359-1399.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational linguistics*, *14*(2), 15-28.

Mourelatos, A. P. (1978). Events, processes, and states. *Linguistics and philosophy*, *2*(3), 415-434.

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge: Cambridge University Press.

Smith, C. S. (1997). *The parameter of aspect*. doi: 10.1007/978-94-011-5606-6

Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca: Cornell University Press.

Verkuyl, H. J. (1972). *On the compositional nature of the aspect*. Springer-Science and Business Media.

Verkuyl, H. J. (1993). *A theory of aspectuality: The interaction between temporal and atemproal structure*. Cambridge: Cambridge University Press.

Verkuyl, H. J. (2005). Aspectual composition: Surveying the ingredients. In H. Verkuyl, H. De-Swart & A. Van-Houtin (Eds.), *Perspectives on aspects* (pp. 19-41). Springer.

Abbreviations

The abbreviations used in the present paper are as follows:

Pl: plural
Sg: singular
Imp: imperfective
pst: past
prs: present
pp.:   past participle
per: perfect
Ez: Ezafe
def: definiteness
indef: indefiniteness
lv: light verb
Pro: progressive
S/A: specificity and accusative case (as indicated by rā)

EMPTY CATEGORY IN PERSIAN RELATIVE CLAUSES

Mahmoud MOBARAKI
Jahrom University, Iran
mmobaraki@jahromu.ac.ir

Abolfazl MOSAFFA JAHROMI
Jahrom University, Iran
a.mosaffa.jahromi@gmail.com

Abstract

Empty categories are one of the fundamental parts of generativist's view towards language. Empty operators, the so-called null elements, which are syntactically active in relative clauses and possibly move into [spec, CP], have reportedly been found in different languages. However, there is no solid evidence for the existence of empty operators in Persian relative clauses. Despite this, syntactic evidences such as theta theory, argument structure and subjacency condition provide satisfactory provable tests in favor of their role in Persian grammar. Namely, Persian relative clauses contain resumptive pronouns which may be covert. Their movement into [spec, CP] can precisely account for subjacency effect in relative clauses. Resumptive pronouns can occupy the subject as well as the object position. This articles attempts to introduce empty operators to Persian syntax, and shows that in such a way it is possible to account for the peculiar behavior of the Persian complementizer "ke" and its obligatory nature.

Keywords: relative clause; resumptive pronoun; empty category; subjacency condition; predicate logic

Povzetek

Ničelne kategorije so eden izmed osnovnih idej generativističnega pogleda na jezik. Ničelni operatorji (tudi ničelni elementi), ki so skladenjsko aktivni v oziralnih odvisnikih in lahko preidejo v [spec, CP], so že bili dokazani v številnih jezikih. Njihova uporabnost pa do sedaj še ni bila pokazana za oziralne odvisnike v perzijščini. Ne glede na to pa sklepamo, da se skladenjski pristopi kot so teorija theta, struktura argumenta in pogoj podstave izkazujejo kot zanesljivi testi, ki spodbujajo uporabo ničelnih operatorjev v perzijski slovnici. Oziralni odvisniki v perzijščini namreč vsebujejo rezumptivne zaimke, ki so lahko prikriti. Njihova sprememba v [spec, CP] lahko natančno pojasni vpliv sintaktične omejitve pojavljanja (angl. Subjacency effect) oziralnih odvisnikih. Razumptivni zaimki namreč lahko zasedejo položaj tako osebka kot tudi predmeta. Članek poskuša uvesti ničelne opratorje v perzijsko skladnjo in pokaže, da lahko na ta način razložimo značilno obnašanje perzijskega dopolnike "ke" ter njegove obvezne pojavnosti.

Ključne besede: oziralni odvisnik; rezumptivni zaimek; ničelna kategorija; pogoj podstave; logika povedka

1   Introduction

One of the issues in government and binding theory is empty categories. These categories can appear as subject or object in clauses in different languages. Soundless categories are NP-trace, Wh-trace, PRO or pro (Haegeman, 1992). Empty categories form the fundamental part of the government and binding model. Although there have been some changes added to the theory in recent years, empty categories have played an important and considerable role in grammar, explaining many grammatical realities (Chomsky, 1995; Radford, 2006; Hornestein et al., 2006). Empty categories are the consequence of projection principle, which says that the structure and lexical information must have syntactic representation at all levels. In other words, it means that categories exist at all levels. The projection of lexical information at the syntactic level can have phonetic representation or not. In second alternation, this information represents an empty category. Although empty categories are phonetically empty, they – according to the trace theory, projection principle and binding theory – exist in mental representation. If the movement does not trace, much of the realities remain unjustifiable in language. Empty categories are the consequence of universal grammar principles, such as the ones which are provable in relative clauses.

There are also empty categories in Persian relative clauses which leave behind a trace. Traces are also under the subjacency principle and empty category principle from which deviation leads to ungrammaticality of the sentence (Riemsdijk & Williams, 1986). The movement of constituents should also be considered in relative clauses. The question which arises here is whether movement happens in Persian relative clauses or not. In case the answer is "yes", the following questions is whether there is an evidence for this movement in Persian relative clauses. Regarding relative clauses, there exist three aproaches in previous findings: "ke: that" movement and empty operator movement. The first approach has been cited in Miremadi (2008, p. 175) in the framework of government and binding theory (GB). Miremadi considers "ke" as a wh-word which moves from its original position to the specifier position of the complementizer phrase (CP). Ahangar (2000) also like Miremadi (2008) accounts "ke" as a relative pronoun which moves obligatorily from its original position inside relative clause (RC) to the [spec, CP] and leaves behind a trace. This approach has been criticized with the explanation that "ke" is not a pronoun and hence it cannot move (Taghvaipour, 2005). The fact that Persian does not include relative pronoun is discussed by Safavi (1994, p. 191), Khayampour (1973, p. 116) also by Lazard (1957, p. 229) and Windfuhr (1978, p. 62).

The other approach for relative clause analysis in Persian considers movement of empty operator. This account is based on the assumption that an empty operator is like a covert relative pronoun and then moves to the [spec, CP]. In this approach, "ke" is defined as a complementizer.

The third approach, cited in Youhanaee (1997, p. 88), provides a view in the framework of the minimalist program. She treats subject and direct object relative clauses by the movement of a null operator, while other types of relative clauses are base generated.

This article tries to show that a movement is also possible in Persian relative clauses, the issue for which up until now not enough linguistic evidence has been found to be either proven or rejected. This article will further explain the nature of this category, and present explanation why "ke" in Persian is obligatory.

2   Materials and methods

To study empty category in Persian relative clauses, the authors tried to consider on casual sentences which are used in standard Persian, and applied syntactic evidence such as theta theory, argument structure and subjacency condition to find further evidence which support the existence of empty category and its behavior in Persian.

3   Discussion

3.1   Relative clause

From the view of language typology, three types of relative clauses can be defined according to the lexical arrangement. In post-nominal type, a relative clause comes after a noun (head). In prenominal type, a relative clause comes before a noun (head) and in adnominal type a head (noun) comes within a relative clause. Persian belongs to the third type, where a relative clause is used after a noun according to Comrie (1989, p. 139), as in the following example.

(1) pesar-i   [ke   shenā mi-konad]   Farānsavi.
    boy-RES[1] [COMP swim   PRES-do-3sg] French
    'The boy who is swimming is French.'

Persian relative clauses always start with "ke", which is necessary for a sentence to be grammatically correct. Example (2) shows grammatically incorrect sentence.

---

[1] This particle precedes restrictive relative clauses in Persian and is shown, henceforth, by RES in gloss.

(2) *pesar-i [shenā mi-konad]   Farānsavi ast.

boy-RES [swim PRES-do-3sg] French   be-PRES-3sg.

'The boy is swimming is French.'

This characteristic differentiates the usage of a "ke- clause" after a noun from the usage of a "ke- clause" in the object position of some verbs such as "dānestan" (to know), "goftan" (to say) etc. In applying a "ke-clause" as one of the arguments of the verb, "ke" is optional and can be deleted:

(3) mi-dān-am   [ke   mi-āy- ad.]

PRES-know-1sg [COMP PRES-come-3sg.]

'I know that s/he comes.'

(4) mi-dān-am   [mi-āy-ad].

PRES-know-1sg [PRES-come-3sg].

'I know s/he comes.'

Dependent clauses in Persian always come with a "ke" as a complementizer, and are thus called complementizer clauses.

3.2   Forming a relative clause

According to transformational grammar, relative clause is a surface structure representation of an embedded sentence which comes after proper or improper noun and has a coreferential NP with a head. Accordingly, sentence (5) with a deep structure converts into a surface structure sentence (6) by changing "pesar" into a pronoun in a complementizer clause:

(5) pesar-i [ke   man pesar rā   zadam]   Farānsavi boud.

boy-RES [COMP I   boy   object marker beat-PAST-1sg] French   be-PAST-1sg

'The boy that I beat the boy was French.'

(6) pesar-i [ke   man ou rā   zadam]   Farānsavi boud.

boy-RES [COMP I   him object marker beat-PAST-1sg] French   be-PAST-1sg

'*The boy that I beat him was French.'

The pronoun used in a complementizer clause, which is related to the head (noun) of the relative clause, is called a resumptive pronoun (Trask, 1993). A resumptive pronoun can be used in the position of a subject, object or a complement of a preposition; the

main point is that a resumptive pronoun is not obligatory and can be deleted. The other point is that Persian sentences, in which the resumptive pronouns are deleted, are considered as unmarked sentences. The following examples show the possibility of deleting resumptive pronouns, and their marked or unmarked counterparts:

(7) a. pesar-i   [ke   Jim rā   zad.] (unmarked)
       boy-RES [COMP Jim object marker beat-PAST-3sg.]
       'The boy that beat Jim.'

   b. pesar-i   [ke   ou Jim rā   zad.] (marked)
       boy-RES [COMP he Jim object marker beat-PAST-3sg.]
       '*The boy that he beat Jim.'

(8) a. pesar-i   [ke   zad-am]. (unmarked)
       boy-RES [COMP beat-PAST-1sg.]
       'The boy that I beat.'

   b. pesar-i   [ke   ou rā   zad-am]. (marked)
       boy-RES [COMP him object marker beat-PAST-1sg.]
       '*The boy that I beat him.'

In spite of the above fact, if the resumptive pronoun comes after a preposition, it can not be deleted:

(9) pesar-i   [ke   ketāb rā   be ou dād-am.]
    boy-RES [COMP book object marker to him give-PAST-1sg.]
    '*The boy that I gave the book to him.'

(10) *pesar-i [ke   ketāb rā   be dād-am.]
      boy-RES [COMP book object marker to give-PAST-1sg.]
      'The boy that I gave the book to'

The question which arises here is whether, in the case of a deleted resumptive pronoun, the movement of the constituent has taken place or not. The hypothetical answer is that if the absence of a resumptive pronoun is not the consequence of the movement of the constituent, deletion will take place in a relative clause. The question is elaborated in the following parts.

3.3   Resumptive pronoun in subject position

Aresumptive pronouns can be in the subject position as in the following example:

(11) pesar-i   [CP ke   [IP ou deraxt rā   shekaste-ast.]]

    boy-RES [COMP [he   tree   object marker break- PRES P-3sg.]]

    '*The boy that he has broken the tree.'

In Persian, the representation of a resumptive pronoun can be soundless as in the following example:

(12) pesar-i   [CP ke   [IP deraxt rā   shekaste –ast.]]

    boy-RES [COMP [tree   object marker break- PRES P-3sg.]]

    'The boy that has broken the tree.'

The question which can be posed for Persian is whether the absence of phonetic representation of pronouns in relative clauses is the consequence of deletion or the result of the movement of a constituent to the specifier position of the complementizer phrase. This is a fundamental question and should include a reasonable syntactical evidence. If the movement of a resumptive pronoun takes place, it should meet the subjacency condition likewise other movements (Haegeman, 1992). However, presentation of evidence on the movement of a pronoun in the subject position in case of a complementizer clause in Persian is unsubstantiated for two reasons. The first issue is the property of deleting the subject in pro-drop languages such as Persian. In Persian, information of person and number has been coded in a verb, so we can delete the subject without leading to the ungrammaticality of the sentence. So, the absence of a resumptive pronoun in the subject position in Persian can be the result of the nature of being pro-drop rather than its movement. In the following example the deletion of a pronoun takes place in the surface structure, not by moving:

(13) a. Mina goft   [CP ke   [IP ou be madrese mi-ravad.]]

      Mina say-PAST-3sg [COMP [she to school   PRES-go-3sg.]]

      'Mina said that she goes to school.'

    b. Mina goft   [CP ke   [IP be madrese mi-ravad.]]

      Mina say-PAST-3sg [COMP [to   school   PRES-go-3sg.]]

      '*Mina said that goes to school.'

The second issue is that presenting a strong reason for the movement or deletion of a pronoun needs to be confirmed by the subjacency condition in relative clause. The violation of the subjacency condition can be the evidence for moving the constituent. Chomsky (1981, p. 56) states that α movement must obey the subjacency condition. One of the condition in which we can investigate this matter is the wh-movement, which takes place in a relative clause. In this case, the movement of the resumptive pronoun accompanied by with a wh-movement leads to the violation of the subjacency condition and the creation of an ungrammatical sentence. Take the following sentence into consideration:

(14) mi-shenās-am   [NP pesar-I [CP ke   [IP mi-dān-am   [IP che kār mi-kon-ad.]]]]

    PRES-recognize-1sg [boy-RES   [COMP [PRES- know-1sg [what do PRES-do-3sg.]]]]

    '*I recognize the boy that I know what does he do.'

Sentence (14) is grammatical in Persian. Because of moving both the resumptive pronoun and wh-question, the subjacency condition must be obviated and leads to ungrammatical sentence; but the matter is that Persian language has a great tendency to use wh-question in its position without moving the constituent to the first position of the sentence (Miremadi, 2008), as in the following examples:

(15) Mina dar hayāt che   kār mi-konad?

    Mina in   yard   what do PRES-do-3sg.

    'What does Mina do in the yard?'

(16) Mina kojā   ketaxb rā   mi-barad?

    Mina where book   object marker PRES-take-3sg.

    'Where does Mina take the book?'

(17) Mina kojā   che   kār mi-konad?

    Mina where what do PRES-do-3sg.

    '*where does Mina what do?'

In using a resumptive pronoun in the subject position in Persian, examining whether the movement of the resumptive pronoun takes place or not is problematic and there is no sufficient evidence to make conclusions. However, the object position can present some evidence for this matter. The following part deals with the study of resumptive pronouns in the object position.

3.4   Resumptive pronoun in object position

Resumptive pronouns can be used in the object position as in example (18).

(18) pesar-i   [_CP_ ke   [_IP_ Mina ou rā   davat-xāhad kard.]]
     boy-RES [COMP [Mina   him object marker invite-FUTURE-3sg.]]
     '*The boy that Mina will invite him.'

The resumptive pronoun in the object position can be deleted, as the following example shows:

(19) pesar-i   [_CP_ ke   [_IP_ Mina davat-xāhad kard.]]
     boy-RES [COMP [Mina   invite-FUTURE-3sg.]]
     'The boy that Mina will invite.'

In the following sentence (20), we use "**t**" in the position of the resumptive pronoun; that means we hypothesize that the movement has taken place.

(20) mi-shenās-am   [_NP_ pesar-i [_CP_ ke   [_IP_ Mina goft   [_IP_ Jim **t** fardā
     PRES-know-1sg [boy-RES [COMP [Mina   say-PAST-3sg [Jim   tomorrow

     davat -xāhad kard.]]]]
     invite-FUTURE-3sg.]]]]

     '*I know the boy that Mina said Jim will invite tomorrow.'

In the above sentence, [_CP_ ke [_IP_ Mina goft]] is considered as a barrier and the trace can not be coreferential with the head "pesar". It means that the reference of the trace is in the domain of second IP. Therefore, it is expected that the second constituent would not move because this would lead to the obviation of subjacency condition and further the ungrammaticality of the sentence. This can be shown empirically by writing a wh-question instead of "fardā" (tomorrow):

(21) mi-shenās-am   [_NP_ pesar-i [_CP_ ke   [_IP_ Mina goft   [_IP_ Jim **t** che moqe
     PRES-know-1sg [boy-RES [COMP [Mina   say-PAST-3sg [Jim   when

     davat -xāhad kard.]]]]
     invite-FUTURE-3sg.]]]]

     '*I know the boy that Mina said when will Jim invite.'

In example (20) "fardā: tomorrow" is in the unmarked position, which is the meaning in its deep structure position. In case "che moqe: when" is moved from its position to before "Jim", we encounter with two cases. If we have another movement before the movement of "che moqe", the movement of the second constituent leads to the obviation of the subjacency condition and hence to the ungrammaticality of the sentence. On the other hand, if the movement takes place before the movement of "che moqe", the sentence is grammatical. The ungrammaticality of the following sentence leads us to conclude that before the movement of "che moqe" we had had another movement, unless the ungrammaticality of the following sentence is not justifiable.

(22) *mi-shenās-am [NP pesar-i [CP ke   [IP Mina goft   [IP che moqe Jim **t₁ t₂**

   PRES-know-1sg [boy-RES [COMP [Mina   say-PAST-3sg [when   Jim

   davat -xāhad kard.]]]]

   invite-FUTURE-3sg.]]]]

   '*I know the boy that Mina said when will Jim invite.'

We may therefore conclude that the ungrammaticality of the above sentence is justifiable when we believe in a presence of some other movement before the movement of "che moqe". In the above clause, the movement of the resumptive pronoun had taken place and because of the movement of the second constituent, the subjacency condition is obviated; therefore, we expect that in the case of inserting the resumptive pronoun, the above sentence becomes grammatical. This is demonstrated by the following example.

(23) mi-shenās-am   [NP pesar-i [CP ke   [IP Mina goft   [IP che moqe   Jim ou

   PRES-know-1sg   [boy-RES [COMP [Mina   say-PAST-3sg [when   Jim him

   rā **t₂**   davat -xāhad kard.]]]]

   object marker   invite-FUTURE-3sg.]]]]

   '*I know the boy that Mina said when will Jim invite him.'

At this point we can conclude that if we have a movement in the complementizer clause in which the resumptive pronoun in the object position does not have a phonetic representation in s-structure, it is reasonable that we have a movement in the subject position. Consequently, relative clauses may also represent the cause of the movement. Another question which arises here is whether there is other evidence to support this claim. Theta theory presents other evidence for approving the movement and the existence of an empty category in Persian relative clauses.

3.5   Theta theory and the movement in relative clauses

Theta theory in government and binding theory deals with the valency of the verbs (Trask, 1993). Determining the necessary constituents of a sentence through predicate logic is a part of theta theory (Haegeman, 1992) and this is the lexical information which determines the number of arguments of a verb (Cowper, 1996). Verbs which take an object are two place predicates (Hurford & Heasly, 1996). The verb "davat kardan: to invite" is a two place predicate, which gives the theta role "patient" to the object position as in the following example:

(24) Mina ou   rā   davat-xāhad kard.

Mina him/her object marker invite-FUTURE-3sg.

'Mina will invite him/her.'

Deletion of one of the arguments leads to the ungrammaticality of the sentence:

(25) *Mina davat-xāhad kard.

Mina   invite-FUTURE-3sg.

'*Mina will invite.'

Let us consider the case (26), in which the resumptive pronoun does not have a phonetic representation in the object position; if we believe in deletion, we will encounter a problem.

(26) mi-shenas-am   pesar-i   [$_{CP}$ ke   [$_{IP}$ Mina davat-xāhad kard.]]

PRES-know-1sg boy-RES [COMP [Mina   invite-FUTURE-3sg.]]

'I know the boy that Mina will invite.'

The object position of the verb "davat kardan: to invite" is governed and receives a case. The position which receives a case must have a phonetic representation (Cowper, 1996). In case of believing in deletion, the grammar wrongly predicts that the above sentence must be ungrammatical, and so its grammaticality is unjustifiable. But if we believe in the movement and the substitution of the constituent, the absence of a phonetic representation is not problematic because the moved constituent settles in the position with no case. Deletion of an object leads to the change of meaning, which in other words means that the transformation changes the meaning. This goes against the Katz-Postal (1964) hypothesis, which says that transformations preserve the original meanings. At this point it is therefore necessary to believe in the movement of a constituent in relative clauses where pronouns in the object position are soundless.

In the object position, we can intervene between the head and the complementizer clause by inserting other clauses without making a sentence ungrammatical, such as in the following example.

(27) mi-shenas-am   pesar-i   [_{CP} ke   [_{IP} fekr mi-konad [_{IP} Mina xāb-dide- ast

     PRES-know-1sg boy-RES [COMP [think-PRES-3sg   [Mina   dream-PAST-3sg

     [_{IP} hame **t**   doust-dārand   [_{IP} ou rā   bebinand.]]]]]

     [all   like-PRES-3pl [her object marker   see-PRES-3pl.]]]]]

     '*I know the boy that he thinks Mina had dream all like to see her.'

Example (27) shows that we can not consider the meaning of the resumptive pronoun (in doust dāshtan clause) in relation to "pesar-i" (head). Instead, the complementizer clause acts independently from the head both from the point of meaning and the point of syntax. The object of the verb "doust dāshtan", which is the soundless resumptive pronoun, must thus exist to project but the grammar nevertheless wrongly predicts the above sentence as ungrammatical. According to the above claims, we may conclude that rather than deletion it is the movement that takes place in relative clauses. If a relative clause without a resumptive pronoun is the cause of the movement, the moved constituent must leave behind a trace. The following part discusses the nature of this trace.

3.6   The nature of the trace of the resumptive pronoun movement

The authors have tried to prove the existence of an empty category in Persian relative clauses so far, and used the predicate logic, theta theory, subjacency condition, and Katz-Postal hypothesis to prove that the movement of a resumptive pronoun takes place in relative clauses. Since the movement of the constituents leaves behind a trace, it is necessary to study the nature of such a trace.

The trace of a resumptive pronoun in Persian clauses is considered as one of the empty categories. The idea of the existance of empty categories is one of the characteristics of the government and binding theory, and it refers to the position that is active syntactically but it does not have any phonetic representations (Cowper, 1996). Four types of empty categories are differentiated in this theory: NP- trace, Wh-trace, pro and PRO. These empty categories have different distributions in sentences according to different principles of universal grammar (Horrocks, 1987) and they always occupy the position of a noun phrase in a sentence. Several parts of the government and binding theory such as the control theory, government theory, and binding theory are related to empty categories (Trask, 1993). Consider the following sentence with a relative clause:

(28) mi-shenās-am   [_{NP} pesar-i [_{CP} ke   [_{IP} Mina **e** davat-xāhad kard.]]]

PRES-know-1sg [boy-RES [COMP [Mina   invite-FUTURE-3sg.]]]

'I know the boy that Mina will invite.'

The empty category "**e**" is in the position where it is governed and receives a case. Consequently, it is not possible for "**e**" to be PRO because PRO can not be governed (Lasnik, 1988). The empty category "**e**" can not be the trace of NP, because the trace of NP does not receive case (Lasnik, 1988). As for other empty categories, pro and wh-trace are left to be evaluated. The category of pro receives case and is governed (Carnie, 2002) and just the principle B from the government and binding theory is applicable to it (Chomsky, 1982, p. 78). The "**e**" category is in the position which is governed and receives case, so it is expected to be pro. However, if "**e**" is perceived as pro, it is not limited by subjacency condition because pro is not the result of a movement. If "**e**" is the same as pro, the ungrammaticality of the following sentence will be unjustifiable:

(29) *mi-shenās-am [_{NP} pesar-i [_{CP} ke   [_{IP} Mina goft   [_{IP} che moqe Jim **e**

PRES-know-1sg [boy-RES [COMP [Mina   say-PAST-3sg [when   Jim

davat-xāhad kard.]]]

invite-FUTURE-3sg.]]]

'*I know the boy that Mina said when will Jim invite.'

The above sentence is ungrammatical when a movement had taken place before wh-movement in a relative clause. Pro is not the result of a movement and the ungrammaticality of the above sentence is the witness against the matter that "**e**" can not be pro.

The only option remaining is the wh-trace, the position from which the wh-question moves and receives case. When a wh-question moves, it must obey the subjacency condition. Ungrammaticality that rises from such a situation is demonstrated by the following sentence:

(30) *mi-shenās-am [_{NP} pesar-i [_{CP} ke   [_{IP} Mina goft   [_{IP} che moqe Jim **e t**

PRES-know-1sg [boy-RES [COMP [Mina   say-PAST-3sg [when   Jim

davat xāhad kard.]]]

invite-FUTURE-3sg.]]]

'*I know the boy that Mina said when will Jim invite.'

The ungrammaticality of the above sentence is justifiable when we believe that another wh-movement had taken place before the movement of "che moqe", leaving behind a trace. In other words, the position in which "che moqe" wants to settle had been occupied before. If this claim is true, inserting a resumptive pronoun leads to grammaticality of the sentence as the following sentence shows:

(31) mi-shenās-am   [NP pesar-i [CP ke   [IP Mina goft   [IP che moqe Jim ou
     PRES-know-1sg [boy-RES   [COMP [Mina   say-PAST-3sg [when   Jim him

     rā   **t**   davat -xāhad kard.]]]
     object marker   invite-FUTURE-3sg.]]]

     '*I know the boy that Mina said when will Jim invite him.'

So, there must be a soundless reference to "**e**" in Persian relative clauses without a resumptive pronoun. This element, which is shown by "**o**" in texts, has been produced in a D-structure as the complement of the verb "davat kardan", and has moved to [spec, CP] and left behind a trace which is shown by "**e**". Thus, "**e**" is the trace of a wh-question. The D-structure and S-structure of the complementizer clause are presented respectively in the following:

(32) [NP pesar-i   [CP ke   [IP Mina goft   [IP Jim **o** davat- xāhad kard]]]
     [boy-RES   [COMP [Mina   say-PAST-3sg [Jim   invite-FUTURE-3sg]]]

     rā   mi-shenās-am.
     object marker PRES-know-1sg.

     '*I know the boy that Mina said Jim will invite.'

(33) [NP pesar-i   [CP **o** ke [IP Mina goft   [IP Jim **e** davat- xāhad kard]]]
     [boy-RES   [COMP [Mina   say-PAST-3sg [Jim   invite-FUTURE-3sg]]]

     rā   mi-shenās-am.
     object marker PRES-know-1sg.

     '*I know the boy that Mina said Jim will invite.'

"o" is one of the empty categories (Haegeman, 1992) that has not been analyzed so far in Persian. This element is called an empty operator. The question which arises here is that what kind of empty category "**o**" is. "**o**" is not a trace, because it exists in a D-structure. The nature of this category is not so specified, but Jaeggli (1981) considers "**o**" as being the same as PRO. PRO settles in the position without government (Chomsky, 1981, p. 74), "**o**" is in the position which governs. The condition of not having

government for PRO is not one of the principles of the theory but the consequence of the PRO characteristic ([+ pronoun, +anaphor]). Therefore, in a D-structure PRO can settle in a position that governs, and since the government and binding theory show its function in S-structure, its occurrence in the position where government is donated is not problematic (Chomsky, 1981, p. 74). In a S-structure PRO moves to the position without government, it means to the specifier position of the complementizer clause.

3.7   Movement and the principle of the empty category

Above we have shown that there is a soundless category "**o**" in Persian relative clauses. "**o**" is one of the arguments of a verb in a D-structure which is transferred to [spec, CP] in a S-structure. Jaeggli (1981) considers "**o**" as a PRO, which necessarily makes it move to the position without government in a S-structure. We can thus conclude that CP is a barrier for government, as in the following sentence.

(34) [$_{NP}$ pesar-i [$_{CP}$ **o** ke [$_{IP}$ Jim **e** davat-xāhad kard]]]   rā   mi-shenās-am.

　　[boy-RES [COMP [Jim   invite-FUTURE-3sg]]] object marker PRES-know-1sg.

　　'I know the boy that Jim will invite.'

"**o**" can not govern its trace. The existence of a complementizer "ke" and its characteristic of being a barrier prevents "**o**" to govern its trace. According to the empty category principle, traces must be governed (Culicover, 1997; Lasnik, 1988), and following it would make the above sentence ungrammatical. To observe the empty category principle and hence the grammaticality of the above sentence, Pesetsky (1982, p. 306) presents a rule for this structure that converts the empty category "**o**" and the complementizer to a constituent which has all of the characteristics of the empty category. In this case, the trace is properly governed. This shows that the complementizer of a relative clause is not a usual complementizer, but it acts as a resumptive pronoun. It is the first time that the issue, which offers a justification for some of the linguistic realities in Persian, has gained attention. Again, the fact is that a complementizer can not be deleted in a relative clause, and its deletion leads the sentence to be ungrammatical:

(35) [$_{CP}$ pesar-i ke   Jim davat -xāhad kard] doust-e-man   ast.

　　[boy-RES COMP Jim invite-FUTURE-3sg] friend-of-mine be-PRES-3sg.

　　'The boy that Jim will invite is my friend.'

(36) *[CP pesar-i Jim davat-xāhad kard]   doust-e man   ast.

[boy-RES   Jim invite-FUTURE-3sg] friend-of-mine be-PRES-3sg.

'The boy Jim will invite is my friend.'

The above issue is against the nature of a complementizer in the following sentences where a complementizer clause is used in a position of one of the arguments of a verb. In this case the complementizer "ke" is optional and can be deleted without leading to the ungrammaticality of the sentence as the following examples show:

(37) shenide-am   [CP ke   Jim pesar-i   rā   davat-xāhad kard.]

hear-PAST-1sg [COMP Jim boy-IND object marker invite-FUTURE-3sg.]

'I heard that Jim will invite a boy.'

(38) shenide-am   [CP Jim pesar-i   rā   davat xāhad kard.]

hear-PAST-1sg [Jim   boy-IND object marker invite-FUTURE-3sg.]

'I heard Jim will invite a boy.'

3.8   Complementizer in relative clause

The above sections have shown that wh-movement also takes place in Persian relative clauses without pronoun, and that the soundless category "**o**" in a D-structure exists in the governor position. Forming a relative clause does not necessarily need a movement in Persian, which is demonstrated by the following example.

(39) mi-shenās-am   [NP pesar-i [CP ke   [IP ou rā   davat -karde-i.]]]

PRES-know-1sg [boy-RES   [COMP [him object marker invite-PRESENT P-2sg.]]]

'*I know the boy that you have invited him.'

In the above sentence, the resumptive pronoun exists in its first position in a D-structure. The result is that the movement has not taken place in this structure, and therefore the complementizer "ke" is expected to have the possibility of being deleted. However, despite our expectation the deletion of the complementizer leads to the ungrammaticality of the sentence:

(40) *mi-shenās-am [NP pesar-i [CP [IP ou rā   davat-karde-i.]]]

PRES-know-1sg [boy-RES   [[him object marker invite-PRESENT P-2sg.]]]

'*I know the boy you have invited him.'

A justification for this behavior can be found in diachronic changes. Several studies have tried to find the origin of a relative clause and its positions in the Indio-European mother language (Harris & Campbel, 1995). Kent (1953) believed that relative clauses in Old Persian were of different types, but most of them had the role of a limiter. In Old Persian, relative, question, complement and conjunction had different forms. According to Abolqasemi (1996, p. 358) complementizers "hya" and "yox" were placed after a noun and could have agreement with a noun in number, case and gender.

(41) **Old Persian:**

    Baga vazrka Ahuramazdā hya imaxm buxmixm adax

    God Big   Ahuramazda who this   universe create-PAST-3sg

    'Ahuramazda is the big God who created this universe.'

In the above example "hya" has agreement with the noun "Ahuramazda" in agentive case.

(42) **Old Persian:**

    axat hex mraom   aẓm yox ahurox mazdax.

    Then him tell-PAST-1sg I   that Ahuramazda.

    'Then, I told him that I am Ahuramazda.'

The complementizer "yox" has agreement with the pronoun "aẓm" in agentive case, gender (male) and number (singular).

In Old Persian, the complementizer had different types of cases such as agentive, accusative and genitive. It also conjugated for singular, dual, plural, masculine and feminine (Abolqasemi, 2016, p. 19). This issue means that the complementizer in a relative clause is not only a complementizer in Persian. Even though surface representations in the Persian language system are missing, the roles have been kept. Thus having case for a complementizer in Persian offers a justification for its impossibility to be deleted. Still, this matter needs a deeper investigation.

In Middle Persian the relative and question forms were combined to "ix", but the complement and conjunction still had different forms (Abolqasemi, 2016, p. 19); In Modern Persian "ke: that" is replaced for all the above forms (Estaji, 2000) and in Modern Persian we consider "ke" a complemenizer of a relative clause.

4   Conclusion

Regarding relative clauses, the article mentioned three approaches. The first approach considers "ke" as a wh-word which moves from its original position to the specifier position of the complementizer phrase. The second approach considers the movement of an empty operator to [spec, CP]. The third approach provides a view in the framework of the Minimalist program. This study followed empty operator approach in the framework of GB. The Persian relative clauses can be produced either with or without a resumptive pronoun. In the usage of the relative clause in which the resumptive pronoun is deleted, the soundless empty category is syntactically active and moves to [spec, CP]. Nevertheless, we do not have a complete justification for the soundless empty category "o" in Persian. This may due to the Persian language and its system, which can not present enough evidence for proving the existence of the soundless empty category and the movement.

This article showed the existence of an empty category in Persian relative clauses. It further presented a justification for "ke" being obligatory in Persian relative clauses, and case in Old Persian was presented. We used the predicate logic, subjacency condition and diachronical view to prove this category.

References

Abolqasemi, M. (2016). *A Historical Grammar of Persian Language*. Tehran: Samt.

Abolqasemi, M. (1996). *The History of Persian Language*. Tehran: Samt.

Ahangar, A. A. (2000). *The Subordinate Clause Construction in Persian based on Government and Binding Theory*. Ph.D. Dessertation. Tehran University.

Carnie, A. (2002). *Syntax: a generative introduction.* Chichester, West Sussex, UK: Wiley-Blackwell.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding.* Cambridge, Mass: MIT Press.

Chomsky, N. (1981). *Lectures on government and binding.* Studies in generative grammar. Dordrecht: Foris.

Chomsky, N. (1995). *The minimalist program.* Cambridge, Mass: MIT Press.

Comrie, B. (1989). *Language universals and linguistics typology*. Oxford: Basil Blackwell.

Cowper, E. (1992). *A concies introduction to syntactic theory.* Chicago, London: Chicago University Press.

Culicover, P. W. (1997). *Principles and parameters: an introduction to syntactic theory*. Oxford: Oxford University Press.

Estaji, A. (2000). Historical study of "ke" multifunctions in Modern Persian. *Journal of Linguistics and Khorasan Dialects.*

Haegeman., L. (1992). Introduction to government and binding theory. Oxford: Blackwell.

Harris, A. C., & Campbel, L. (1995). *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Hornstein, H., Nunes, J., & Grohmann, K. (2006). *Understanding minimalist.* Cambridge: Cambridge University Press.

Horrocks, G. (1987). *Generative grammar.* London: Longman.

Hurford, J., & Heasly, B. (1996). *Semantics: a course book.* Cambridge: Cambridge University Press.

Jaeggli, O. (1981). *Topics in Romance syntax.* Dordrecht: Foris.

Katz, J. J., & Postal, P. M. (1964). *An integrated theory of linguistic description.* Cambridge,. Mass: MIT Press.

Kayne, R. (1984). *Connectedness and binary branching.* Dordrecht: Foris.

Kent, R. G. (1953). *Old Persian*. New Haven: American oriental society.

Khayampour, A. (1973). *Persian Grammar*. Tabriz: Tehran Bookshop.

Lasnik, H. (1988). A course in GB syntax: lectures on binding and empty categories. Cambridge, Mass.: MIT Press.

Lazard, G. (1957). *Grammaire du persan contemporain*. Paris: Klincksieck.

Miremadi, S. A. (2008). *Persian Syntax, A G. B-Based Analysis*. Tehran: Samt.

Pesetsky, D. (1982). *Paths and Categories.* Doctoral Dissertation. MIT Press.

Radford, A. (2006). *Minimalist syntax: exploring the structure of English.* Cambridge: Cambridge University Press.

Riemsdijk, H., & Tushyeh, H. (1998). A contrastive analysis of English and Arabic in relativization. *Papers and studies in contrastive linguistics,* 34, 141-152.

Riemsdijk, H., & Edwin, W. (1986). *Introduction to the theory of grammar.* Cambridge, MA: MIT Press.

Safavi, K., (1994). Some Properties of Persian Relative Clause. *Proceedings of the 2nd Conference of Persian Linguistics.* Allameh Tabatabaie University, Tehran.

Taghvaipour, M. A. (2005). *Persian Relative Clause in Head-driven Phrase Structure Grammar*. Ph.D. Desertation. University of Essex.

Trask, R. L. (1993). *A dictionary of grammatical terms in linguistics*. London and New York: Routledge.

Windfuhr, G. L. (1978). *Persian Grammar: History and State of its Study*. The Hague: Mouton Publishers.

Youhanaee, M. (1997). The Acquisition of the English Complementizer Phrase by Adult Persian Speakers. Ph.D. Dissertation. University of Essex.

A Comparative Study on Speculation Forms of Interrogative Sentences in Korean and Japanese

MOON Changhak

Hankuk University of Foreign Studies, South Korea

kudosmoon@gmail.com

Abstract

This article reveals the interrogative aspect of speculation forms in Japanese and Korean, including the characteristics of *-keyss-nya* in Korean and *darooka* in Japanese. *-keyss-nya* and *darooka* by definition indicate meanings of both "speculation" and "interrogation", and thus it is anticipated that the semantic characteristics of these forms will not differ significantly. However, one perceives many differences when examining the semantic characteristics of both. First, a characteristic of *-keyss-nya* is its <listener-oriented interrogative sentences> with strong communicativity, whereas a characteristic of *darooka* is its <speaker-oriented interrogative sentences> with weak communicativity. Second, based on this characteristic of <listener-oriented interrogative sentences>, *-keyss-nya* is not made into "question usage that does not solicit an answer" or "exclamation usage," but it may be made into "strong rhetorical questions." Meanwhile, based on the characteristic of <speaker-oriented interrogative questions> with weak communicativity, *darooka* may be derived into "question usage that does not solicit an answer", "exclamation usage", and "weak rhetorical interrogative sentences." As stated above, different semantic meanings of *-keyss-nya* and *darooka* result from semantic differences between *-keyss-* and *daroo*. The *-keyss-* meaning of "completion of judgement formation" is reflected in *-keyss-nya*'s characteristics of <listener-oriented interrogative sentences>*,* and *daroo*'s meaning of the "judgement formation process" is reflected in *darooka*'s characteristics of <speaker-oriented interrogative sentences>.

Keywords: speculation; interrogation; *-keyss-nya*; *darooka*; judgement formation

Povzetek

Članek razkriva vprašalne vidike domnevnih oblik v japonščini in korejščini, vključujoč značilnosti izrazov *-keyss-nya* v korejščini in *darooka* v japonščini. Po definiciji *-keyss-nya* in *darooka* nosita pomena »domneva« in »povpraševanje«, zato predvidevamo, da se pomenskost teh dveh različih ne bo bistveno razlikovala. Vendar pri obravnavanju pomenskih značilnosti opažamo veliko razlik. Ena le-teh je, da so za korejski *-keyss-nya* značilni vprašalni stavki, ki so usmerjeni proti poslušalcu in imajo močno komunikacijsko vlogo, medtem ko se *darooka* pojavlja v vprašalnih stavkih, ki so usmerjeni h govorcu in imajo šibko komunikacijsko vrednost. Na osnovi teh značilnosti se korejski *-keyss-nya* ne pojavlja v rabi vprašanja, ki ne

zahteva odgovora, ali pa v rabi vzklika, lahko pa deluje kot močno retorično vprašanje. Po drugi strani se japonski *darooka* pojavlja tako v rabi vprašanja, ki ne zahteva odgovora, kot tudi v rabi vzklika ali kot šibko retorično vprašanje. Kot že omenjeno, razlika med *-keyss-nya* in *darooka* izvira iz pomenskih razlik med *-keyss-* and *daroo*. *-keyss-* pomensko nakazuje na dokončnost oblikovanja presoje in se odraža v vprašalnih stavkih, ki so usmerjeni proti poslušalcu, medtem ko pomen izraza *daroo* nakazuje na proces oblikovanja presoje in se pojavlja v vprašalnih stavkih, usmerjenih h govorcu.

Ključne besede: domneva; povpraševanje; *-keyss-nya*; *darooka*; oblikovanje presoje

1   Introduction

Korean and Japanese are known to have numerous typological commonalities. As shown below, one can observe similar aspects of both languages, including the fact that both are agglutinative and their sequences of affixes demonstrate grammatical categories in predicates. [1,2]

(1) [[[[[[*ilk*]   *hi*]   *ko iss*]   *ci anh*]   *ass*]   ***keyss***]   *ta*]
    [[[[[[verb stem]   voice]   aspect]   polarity]   tense]   epist. modality]   commun. modality]
    '(This book) must not have been read.'

(1') [[[[[[[*yom*]   *are*]   *tei*]   *nakat*]   *ta*]   ***daroo***]   *ne*]
    [[[[[[verb stem]   voice]   aspect]   polarity]   tense]   epist. modality]   commun. modality]
    '(This book) must not have been read.'

Among the affixes that demonstrate grammatical categories, *-keyss-* in Korean and *daroo* in Japanese have been shown to commonly demonstrate epistemic modality and speculative meaning as displayed below.[3]

(2) *Nayil   pi-ka   o-**keyss**-ta.*
    Tommorow rain-NM come-SPEC-DECL
    'It may rain tomorrow.'

(2') *Ashita   Ame-ga huru-**daroo**.*
    Tomorrow rain-NM come-SPEC/DECL
    'It may rain tomorrow.'

---

[1] See Horie and Taira (2002), Moon (2015a, 2015b), Nitta (1997), and Umeda (1990) for more details.

[2] Example sentences with no reference were made by the author.

[3] Korean's *-keyss-,* besides its "speculation" meaning, also demonstrates a meaning of "volition" based on the context. The present article limits *-keyss-* to its demonstration of meanings of "speculation" with exception to its meaning of volition.

As shown above, because both -*keyss*- in Korean and *daroo* in Japanese exhibit the meaning of "speculation", they can be understood as forms that semantically correspond with one another. -*keyss*- in Korean and *daroo* in Japanese will be explored in more detail in section 2, including their use in interrogative sentences. The interrogative sentences with the speculation forms in Korean and Japanese both exhibit meanings of "speculation" and "interrogation", and thus the semantic characteristics of these forms are not expected to differ significantly. If this is the case, one may question whether in actuality the semantic characteristics of these forms will be similar.

The studies thus far have produced little regarding analysis of this issue. Accordingly, the present article examines the interrogative sentences of speculation forms in Korean and Japanese and makes a comparison of their semantic characteristics. Furthermore, based on these results, this study intends to reconsider the essence of speculation forms in Korean and Japanese from the perspective of "judgement formation."

2   The morphological characteristics of speculation forms and their interrogative sentences in Korean and Japanese

First, this study examines the morphological characteristics of speculation forms in Korean and Japanese. As mentioned below, -*keyss*- in Korean and *daroo* in Japanese possess a commonality in that they form a type of epistemic modality paradigm comprising an opposition between affirmation and speculation due to the existence of these forms. However, while Korean's -*keyss*- demonstrates only speculation meaning, Japanese's *daroo,* in addition to signifying speculation, also exhibits the meanings of declarative sentence and plain style. That is to say, *daroo* in Japanese is a form in which epistemic modality, sentence type, and speech style are being merged. In Korean, sentence type and speech style are being merged in inflectional suffixes (referred to as sentence-final suffixes in Korean linguistics).

(3) [Korean]
   a. *Pakk-ye   pi-ka   o-ko iss-ta/supnita.* [affirmation]
      Outside-LOC rain-NM come-ASP-DECL·SPEECH (plain style/polite style)
      'It's raining outside.'

   b. *Pakk-ye   pi-ka   o-ko iss-**keyss**-ta/supnita.* [speculation]
      outside-LOC rain-NM come-ASP-SPEC-DECL·SPEECH (plain style/polite style)
      'It must be raining outside.'

(4) [Japanese]
    a. *Soto-ni*   *ame-ga* *hut-tei-ru/masu.* [affirmation]
       outside-LOC rain-NM come-ASP-DECL·SPEECH (plain style/polite style)
       'It's raining outside.'

    b. *Soto-ni*   *ame-ga* *hut-teiru-**daroo/desyoo**.* [speculation]
       outside-LOC rain-NM come-ASP-SPEC·DECL·SPEECH (plain style/polite style)
       'It must be raining outside.'

Next, this study verifies the morphological characteristics of interrogative sentences of speculation forms in Korean and Japanese. As shown below, to employ *-keyss-* in Korean and *daroo* in Japanese into interrogative sentences, interrogative sentence-final form is added. However, the nature of the forms added to make an interrogative sentence in Korean and Japanese is slightly different. Interrogative sentences in Korean are made by adding the inflectional suffixes *-nya* or *-supnikka,* and these inflectional suffixes merged interrogative meaning and speech style. Conversely, interrogative sentences in Japanese are typically made by adding the sentence-final interrogative particle *ka*. Along with other sentence-final particles, *ka* is a part of the category of a "communicative modality".

(5) [Korean]
    a. *Pakk-ye*   *pi-ka*   *o-ko iss-**nya**/**supnikka***? [question of affirmation]
       outside-LOC rain-NM come-ASP-INTER·SPEECH (plain style/polite style)
       'Is it raining outside?'

    b. *Pakk-ye*   *pi-ka*   *o-ko iss-**keyss**-**nya**/**supnikka***? [question of speculation]
       outside-LOC rain-NM come-ASP-SPEC-INTER·SPEECH (plain style/polite style)
       'Would it be raining outside?'

(6) [Japanese]
    a. *Soto-ni*   *ame-ga* *hut-tei-ru/masu-**ka***? [question of affirmation]
       outside-LOC rain-NM come-ASP-SPEECH (plain style/polite style)-INTER
       'Is it raining outside?'

    b. *Soto-ni*   *ame-ga* *hut-teiru-**daroo/desyoo**-ka*? [question of speculation]
       outside-LOC rain-NM come-ASP-SPEC·SPEECH(plain style/polite style)-INTER
       'Could it be raining outside?'

Based on (5) and (6), the subject of analysis in this study—speculation forms and interrogative forms in Korean and Japanese—can be divided and organized based on a sentence style. Table 1 presents the interrogative sentences of speculation forms in Korean as *-keyss-nya* and the interrogative sentences of speculation forms in Japanese as *darooka.*

Table 1: Speculation forms and interrogative sentences in Korean and Japanese

	Speculation form		Interrogative sentence of speculation forms	
	Korean	Japanese	Korean	Japanese
Plain style	*Ha-**keyss-ta***	*Suru-**daroo***	*Ha-**keyss-nya***	*Suru-**daroo**-ka*
Polite style	*Ha-**keyss-supnita***	*Suru-**desyoo***	*Ha-**keyss-supnikka***	*Suru-**desyoo**-ka*

3 Semantic characteristics of -*keyss-nya* and *darooka*

As confirmed in (5b) and (6b), speculation forms in Korean and Japanese can be employed in interrogative sentences. This section compares -*keyss-nya* and *darooka* employed typical interrogative sentences to compare their characteristics.

(7) a. *ku-nun   ilhaknyensayng-i-**pnikka***?
    he-TOP   1st year student-COP- INTER·SPEECH
    'Is he a 1st year student?'

  b. *Kare-wa   itinensee-**desu-ka***?
    he-TOP   1st year student-COP·SPEECH-INTER
    'Is he a 1st year student?'(Moriyama, 1992, p. 68)

In a typical interrogative sentence such as (7), according to Adati (1999, p. 12; 2002, p. 175), the sentence has two primary conditions, as shown below.

(8) Two conditions of a typical interrogative sentence
    a. The speaker lacks some information. <Uncertainty condition>
    b. (S)he requests this information from his or her interlocutor. <Inquiry condition>

First, the <uncertainty condition> in typical interrogatives as shown below can be expressed as alternative questions by listing contradictions in the relevant sentences. The fact that alternative questions are possible and can list contradictions in this manner could mean that, from the speaker's point of view, it is uncertain whether the proposed content is true or false.

(9) a. *ku-nun   ilhaknyensayng-i-**pnikka**,*
    he-TOP   1st year student-COP- INTER·SPEECH

    animyen   *ilhaknyensayng-i ani-**pnikka**?*
    or   1st year student-COP·NEG- INTER·SPEECH

    'Is he a 1st year student? Or, isn't he a 1st year student?'
  b. *Kare-wa   itinensee-desu-**ka**,*

he-TOP   1st year student-COP·SPEECH-INTER

aruiwa   *itinensee-dewaarimasen-**ka**?*
or   1st year student-COP·NEG·SPEECH-INTER

'Is he a 1st year student? Or isn't he a 1st year student?' (Moriyama, 1992, p. 68)

Next, the <inquiry conditions> of a typical interrogative sentence, as shown below, confirm that these sentences are inconsistent as the embedded clauses of mental verbs. The fact that these sentences cannot become embedded clauses of mental verbs demonstrates their communicability, and thus ultimately the fact that they request information from the listener.

(10) a. **ku-nun   ilhaknyensayng-i-**pnikka hako sayngkak hay-ss-ta**.*
     he-TOP   1st year students-COP-INTER·SPEECH QUOT think-PAST-DECL
     'I thought, 'Is he a 1st year student?''

   b. **Kare-wa ichinensee-**desu-ka to omo-t-ta**.*
      he-TOP   1st year student-COP·SPEECH-INTER QUOT think-PAST-DECL
      'I thought, 'Is he a 1st year student?''

On the two conditions above, this paper examines what characteristics *-keyss-nya* and *darooka* possess.

First, to verify the <uncertainty condition>, this study investigates whether it is possible to change the corresponding sentences into alternative questions. As shown below, both *-keyss-nya* and *darooka* can be made into alternative questions and fulfill the <uncertainty condition>.

(11) a. *Pakk-ye   pi-ka   o-ko iss-keyss-**nya**,*
     outside-LOC rain-NM come-ASP-SPEC-INTER·SPEECH,

     animyen   *an*   *o-ko iss-keyss-**nya**?*
     or   NEG   come-ASP-SPEC-INTER·SPEECH

     'Would it be raining outside, or not?'

   b. *Soto-ni   ame-ga hut-teiru-**daroo**-ka,*
      outside-LOC rain-NM come-ASP-SPEC·SPEECH-INTER,

      aruiwa   *hut-tei-nai-**daroo**-ka?*
      or   come-ASP-NEG-SPEC·SPEECH-INTER

      'Could it be raining outside, or not?'

Next, to verify the <inquiry condition>, this study investigates whether the corresponding sentences are suitable as embedded clauses of mental verbs. In other words, *-keyss-nya* fulfills the <condition of inquiry>, but *darooka* does not.

(12) a. **Pakk-ye   pi-ka   o-ko iss-**keyss-nya ko sayngkak hay-ss-ta.***
        outside-LOC rain-NM come-ASP-SPEC-INTER·SPEECH QUOT think-PAST-DECL
        'I thought, 'would it be raining outside?''

   b. *Soto-ni   ame-ga   hut-teiru-**daroo-ka to omo-t-ta**.*[4]
        outside-LOC rain-NM come-ASP-SPEC·SPEECH-INTER QUOT think-PAST-DECL
        'I thought, 'would it be raining outside?''

With the above content, *-keyss-nya* and *darooka* can be recognized as possessing the following characteristics:

(13) a. In the case of *-keyss-nya*, the speaker is unable to make a judgement due to insufficient information concerning the proposed content and is requesting the information missing from the proposed content from the listener.

   b. *Darooka* only demonstrates speaker's inability to make a judgement due to insufficient information on the proposed content, but it is not a request to the listener for information.

   c. Accordingly, while *-keyss-nya* could be called an <interrogative sentence dependent on listener information>, or a <listener-oriented interrogative sentence>, *darooka* could be called an <interrogative sentence non-dependent on listener information>, or a <speaker-oriented interrogative sentence>.[5]

4   Question usage that does not solicit an answer

Based on the analysis in section 3, *darooka* would be expected not be used in a conversational situation because it was shown to be <speaker-oriented interrogative sentence>. However, as is shown below, *darooka* can also be used in conversational situations. This study examines characteristics of *darooka's* conversational usage by comparing them with typical interrogative sentences and *-keyss-nya.*

---

[4] Circumstances slightly change when *darooka* becomes its polite form of *desyooka.* Section 4 explains this in detail.

[5] The terms "interrogative sentences dependent on listener information" and "interrogative sentences non-dependent on listener information" are used as they appear in Moriyama (1992, 2000).

(14)   (The speaker enters the classroom, sees *Taroo* and *Jiroo* in the classroom together, is surprised at the sight and speaks to *Taroo*[6])

a.   ??*Taroo san!*   *Jiroo san-ga   nande   koko-ni   ir-u-n-**daroo**?*[7]
Taroo,   Jiroo -NM   why   here-LOC   exist-ATT-NLMZ-SPEC(-INTER)
'*Taroo*! Why would *Jiroo* be here?'

b.   *Taroo san!*   *Jiroo san-ga   nande   koko-ni   iru-**no**?*
Taroo,   Jiroo -NM   why   here-LOC   exist-INTER
'*Taroo*! Why is *Jiroo* here?'

(15)   (The speaker enters the classroom with *Taroo*. They see *Jiroo* in the classroom, both are surprised, and the speaker speaks to *Taroo*)

a.   *Taroo san!*   *Jiroo san-ga   nande   koko-ni   ir-u-n-**daroo**?*
Taroo,   Jiroo -NM   why   here-LOC   exist-ATT-NLMZ-SPEC(-INTER)
'*Taroo*! Why would *Jiroo* be here?'

b.   *??Taroo san!*   *Jiroo san-ga   nande   koko-ni   iru-**no**?*
Taroo,   Jiroo -NM   why   here-LOC   exist-INTER
'*Taroo*! Why is *Jiroo* here?'

c.   ??*Taroo ssi*   *Jiroo ssi-ga   way   yeki-ey   iss-**keyss-nya**?*
Taroo,   Jiroo -NM   why   here-LOC   exist-SPEC-INTER
'*Taroo*! Why is *Jiroo* here?'

The utterance situation of (14) is one in which the listener (*Taroo*) is in a classroom with *Jiroo* and can see "why *Jiroo* is here." In other words, in this situation the speaker can assume that the listener will know about the proposed content for which the speaker has insufficient information. In a situation such as this, where the listener is assumed to have an answer regarding the speaker's question, *darooka* is not suitable, as shown in (14a), but a typical interrogative sentence is suitable, as shown in (14b).

Conversely, in the utterance situation of (15) the speaker and listener (*Taroo*) enter the classroom together; thus, both the speaker and the listener are unable to know "why *Jiroo* is here." That is to say, in this situation it would be difficult for the speaker to assume that the listener knows about the information lacking regarding the proposed content. In situations such as this, where it is difficult to assume that the listener has an answer to the speaker's inquiry, a typical interrogative sentence such as (15b) is not suitable, whereas a sentence with *darooka*, such as in (15a), is suitable.

---

[6] (14) is a normal situation where the speaker and Taroo are close enough so they have no problem asking and answering questions. I plan to provide an even more natural real-life example at the next opportunity to explain this in detail.

[7] The *ka* is often dropped in *darooka* (or *desyooka*) when used in wh-questions.

Based on the above language behavior of *darooka*, it is obvious that typical interrogative sentences and *darooka* possess the following characteristics. Typical interrogative sentences assume that the listener has an answer to the speaker's inquiry, and they have the function to solicit an answer from the listener based on that assumption. Conversely, *darooka* assumes that the listener will not have an answer regarding the speaker's inquiry and, based on that assumption, has the function of not soliciting an answer from the listener.[8]

Whereas, while in situations such as (15a), where it is difficult for the speaker to assume that the listener will know about the proposed information, *darooka* may be used, *-keyss-nya* may not, as shown in (15c).

However, as shown in (16a), *darooka's* polite form, *desyooka*, may also be used in situations where it is assumed that the listener has an answer regarding the speaker's inquiry. Compared to typical interrogative sentences, *desyooka* conveys a softer and politer nuance.[9] This nuance comes from its function of not soliciting an answer, as with *darooka*, and not pressuring the listener.

However, in situations where it is assumed that the listener has an answer to the speaker's inquiry, *darooka's* polite form *desyooka* may be used with an accompanying soft and polite nuance but, as shown in (16b), *-keyss-nya's* polite form *-keyss-supnikka-* can not be used.

(16) (The speaker enters the classroom, is surprised at seeing *Taroo* and *Jiroo* in the classroom together, and speaks to *Taroo*)
   a. *Taroo san!*   *Jiroo san-ga*   *nande*   *koko-ni*   *ir -u-n-**deshyooka**?*
       Taroo,   *Jiroo -NM*   why   here-LOC   exist-ATT-NMLZ-SPEC(-INTER)
       '*Taroo*! Why is *Jiroo* here?'

   b. ??*Taroo ssi*   *Jiroo ssi-ga*   *way*   *yeki-ey*   *iss-**keyss-supnikka**?*
       Taroo   *Jiroo -NM*   why   here-LOC   exist-INTER
       '*Taroo*! Why is *Jiroo* here?'

As seen above, although *darooka* is also used in conversational situations, based on its fundamental nature as a <speaker-oriented interrogative sentence>, it possesses a "soft and polite nuance as an interrogative sentence that does not solicit an answer from the listener." However, *-keyss-nya* differs from *darooka,* it is not implying question usage that does not solicit an answer from the speaker; thus, it is not a <speaker-oriented interrogative sentence>.

---

[8] Refer to Adati (2002) and Nihongo Kizyutsu Bunpoo Kenkyuukai (2003) for more on the function of *darooka* in which it does not solicit an answer from the listener.

[9] Refer to Moriyama (1992), Adati (2002), and Nihongo Kizyutsu Bunpoo Kenkyuukai (2003) for information about the soft and polite nuance of *desyooka*.

5   Exclamation usage

Darooka, as shown in (17a), accompanies degree or frequency interrogatives and may demonstrate the speaker's exclamatory attitude by implying a high degree or frequency.

(17) a. *Watashi-wa   Tanaka san-ni   **donnani** kanshya shi-te ir-u-koto-**daroo**!*
         I-TOP   tanaka -GOAL   how   thankful-ASP-ATT-NMLZ-SPEC(-INTER)
         'How thankful I am to *Tanaka*!' (Nihongo Kijutsu Bunpoo Kenkyuukai, 2003, p. 88)

     b. ??*Na-nun   tanaka ssi-eykey   **elmana**   kamsa ha-ko iss-**keyss-nya**!*
         I-TOP   tanaka -GOAL   how   thankful-ASP-SPEC-INTER
         'How thankful I am to *Tanaka*!'

The exclamation usage of *darooka* such as in (17a) not only demonstrates speaker's inability to make a judgement regarding insufficient information on the proposed content but is also derived from a <speaker-oriented interrogative sentence> that does not request information lacking from the proposed content. In other words, by demonstrating insufficient information regarding the proposed content in *darooka's* <speaker-oriented interrogative sentence> as a degree or frequency interrogative, the exclamation usage implies a high degree or frequency and naturally assumes the meaning of the exclamation.

As shown above, although *darooka* may be employed in exclamation usage, it achieves the nuance of the exclamation based on its nature as a <speaker-oriented interrogative sentence>. However, *-keyss-nya* in (17b) contrary to *darooka* cannot be employed in "exclamation" usage, and thus, does not mark a <speaker-oriented interrogative sentence>.

6   Rhetorical question (interrogative) usage

Both *-keyss-nya* and *darooka*, as shown below, can pose a question on the premise that listener has established an opposing judgement concerning the proposed content and therefore can be employed for ironic rhetorical question usage that verifies that premise.

(18) a. *Mayil cikak   ha-nun   Chelswu-ka   ilehkey ilccik   o-**keyss-nya**?*
         Every day   late-ATT   *Chelswu*-NM   this early   come-SPEC-INTR
         'Will Chelswu who's late everyday come this early?'

b. ***Celtaylo*** *Chelswu-nun* *ilehkey ilccik* *o-ci anh-**nunta**.*
Never   *Chelswu*-TOP   this early   come-NEG-DECL
'Never would Chelswu come this early.'

c. ***Amato*** *Chelswu-nun* *ilehkey ilccik* *o-ci anh-**Ikesi-ta**.*
Perhaps   *Chelswu*-TOP   this early   come-NEG-SPEC-DECL
'Perhaps Chelswu won't come this early.'

(19) a. *Mainiti tikoku* *su-ru* *Chelswu-ga* *konnani hayaku kuru-**daroo-ka**.*
Every day   late-ATT   *Chelswu*-NM   this early   come-SPEC-INTR
'Would Chelswu who is late every day come this early?'

b. ***Zettai*** *Chelswu-wa* *konnani hayaku* *ko-**nai**.*
never   *Chelswu*-TOP this early   come-NEG·DECL
'*Chelswu*-never comes this early.'

c. ***Tabun*** *Chelswu-wa* *konnani hayaku* *ko-nai-**daroo**.*
perhaps   *Chelswu*-TOP this early   come-NEG-SPEC·DECL
'Perhaps Chelswu wouldn't come this early.'

Both *-keyss-nya* and *darooka* can be employed in ironic rhetorical questions, as shown in (18a) and (19a). However, there appears to be a slight difference between them regarding the possibility of an opposing judgement on the proposed content. When rephrasing (18a)'s *-keyss-nya* rhetorical question into a declarative sentence that demonstrates an opposing judgement on the proposed content, both the adverb *amato* in (18c), which exhibits a low possibility, and the adverb *celtaylo* in (18b), which exhibits a high possibility, are suitable when they collocate. Meanwhile, the *darooka* rhetorical question of (19a) is suitable when it collocates with the adverb *tabun*, which exhibits a low possibility, such as in (19c), but is not suitable when it collocates with the adverb *zettai*, which exhibits a high possibility, such as in (19b). In other words, the rhetorical question with *-keyss-nya* could be interpreted as meaning that the opposing judgement has both, a high and a low possibility, concerning the proposed content, or so it is understood as a "strong rhetorical question", but rhetorical question with *darooka* can only be interpreted as meaning that the possibility of the opposing judgement regarding the proposed content is low, or in other words, it is a "weak rhetorical question."

As stated above, the characteristics of the rhetorical questions of *-keyss-nya* and *darooka* seem to be closely related to the fundamental nature of *-keyss-nya* and *darooka.*

That is to say, a *-keyss-nya* rhetorical question, based on its fundamental nature as a <listener-oriented interrogative sentence>, possesses a nuance that strongly asserts an opposing judgement to the listener concerning the proposed content. However, a

darooka rhetorical question, based on its fundamental nature as a <speaker-oriented interrogative sentence>, only goes as far as the speakers thinking by themselves about an opposing judgement regarding the proposed content or lightly expressing this judgement to the listener.

7   Rethinking speculation forms in Korean and Japanese from the perspective of "judgement formation"

This study has so far examined the semantic characteristics of *-keyss-nya* and *darooka.* Since semantically both demonstrate meanings of "speculation" and "inquiry," their semantic characteristics were not expected to differ significantly, but, as shown, there are many areas where the characteristics of *-keyss-nya* and *darooka* are different.

Employment of *-keyss-nya* results in a <listener-oriented interrogative sentences> with strong communicativity, whereas employment of *darooka* results in <speaker-oriented interrogative sentences> with weak communicativity.

In this manner, the fact that the semantic meanings of *-keyss-nya* and *darooka* differ may be because the speculation forms in Korean and Japanese, that is to say, the meanings of *-keyss-* and *daroo*, are fundamentally different. This section once more reconsiders the semantic characteristics of *-keyss-* and *daroo* in terms of their semantic differences from the perspective of Moriyama's (1992, 2000) "judgement formation". According to Moriyama (1992, p. 73), *daroo* demonstrates that "a conclusion has not yet been reached, or that a judgement is currently being formed."

There is no detailed explanation of "judgement formation" in Moriyama's (1992, 2000) research, but this study aims to interpret Moriyama's "judgment formation" in the following way.

(22) a. "Judgement" means that the "speaker makes a conclusion about whether a judgement has actually been established regarding the proposed content, or makes a judgement about whether a proposition is true or false."

   b. There is a process in "judgement" that includes a "beginning-process-conclusion."

Moriyama (2000, p. 63) states that because *daroo* demonstrates that a conclusion has not yet been made in reality, or that it demonstrates a non-reality "process of judgement formation," *darooka* becomes an <interrogative sentence non-dependent on listener information>. This study can re-summarize Moriyama's thinking as because *daroo* in a declarative is the proposed content for which the speaker's judgement has not been concluded, eventually *darooka* comes back to the speaker even if the proposed content is turned into an interrogative.

This study think that this is so because *daroo* demonstrates a "judgement formation process" in the same way as *darooka* can be derived into speaker-oriented usages such as "question usage that does not solicit an answer," "exclamation usage," and "weak rhetorical questions."

In Moriyama's (1992, 2000) perspective of "judgement formation," this study analyzes *-keyss-* and *-keyss-nya-* as follows: Because *-keyss-* draws a realistic conclusion or, in other words, exhibits a realistic judgment conclusion, it becomes an <interrogative sentence dependent on listener information>. That is to say, because *-keyss-* signals propositional content where the speaker's judgement has been concluded, if one were to make the propositional content into an interrogative, this interrogative could be understood as requesting judgement from the listener with a proposition in which the listener has also completed judgement. In this manner, *-keyss-nya* cannot be made into a "speaker oriented" "question usage that does not solicit an answer" or "exclamation usage," and the fact that it can be used to form a "listener oriented" "strong rhetorical question" is also because *-keyss-* demonstrates a completion of judgement formation.

8   Conclusion

The present article has analyzed interrogative sentences of speculation forms in Korean and Japanese. Results of the analysis are as follows.

(i) To employ speculation forms in interrogations, an inflectional suffix is added to *-keyss-* in Korean that merges speech style and sentence type. In Japanese, however, a particle is added to *daroo* to signal interrogative "communicative modality."

(ii) Korean's *-keyss-nya* satisfies both typical interrogative <uncertainty conditions> and <inquiry conditions>, and sentences employing it possess characteristics of <listener-oriented interrogative sentences> with strong communicativity. *Darooka* in Japanese only satisfies the <uncertainty condition> of typical interrogative sentences and sentences employing it possesse characteristics of <speaker-oriented interrogative sentences> with weak communicativity.

(iii) Korean's *-keyss-nya* cannot be used to produce "question usage that does not solicit an answer" or "exclamation usage" because of its characteristics of <listener-oriented interrogative sentences>, but can be made into a "strong rhetorical question."

(iv) *Darooka* in Japanese can be used to form "question usage that does not solicit an answer," "exclamation usage," and "weak rhetorical questions" because of its characteristic employment in <speaker-oriented interrogative sentences>.

(v) The differences between the semantic characteristics of *-keyss-nya* and *darooka* may be because they semantically mirror opposites. In other words, use of *-keyss-nya*

results in a <listener-oriented interrogative sentence> because -*keyss*- demonstrates "completion of judgement formation," and use of *darooka* results in a <speaker-oriented interrogative sentence> because *daroo* demonstrates the "process of judgement formation."

References

Adati, T. (1999). *Nihongo gimonbun ni okeru handan no shosyoo* [Aspects of judgment in Japanese interrogative clauses]. Tokyo: Kurosio Syuppan.

Adati, T. (2002). Situmon to utagai [Questions and doubts]. In M. Kazuhito et al. (Eds.), *Modariti* [Modality] (Sin Nihongo Bunpoo Sensyo 4, pp. 174–202). Tokyo: Kurosio Syuppan.

Horie, K., & Taira, K. (2002). Where Korean and Japanese differ: Modality vs. discourse modality. In N. Akatsuka & S. Strauss (Eds.), *Japanese/Korean linguistics* (pp. 10, 178–191). Stanford: CSLI.

Moriyama, T. (1992). Nihongo ni okeru suiryoo o megutte [On conjecture in Japanese]. *Gengo Kenkyuu, 101*, 64-83.

Moriyama, T. (2000). Kihon zyohoo to sentaku kankei tosite no modariti [Basic mood and modality as a relation of selection]. In T. Moriyama, Y. Nitta, & H. Kudou (Eds.), *Modariti* [Modality] (Nihongo no bunpoo 3, pp. 1–78). Tokyo: Iwanami Syoten.

Moon, C. (2015a). A contrastive study of Korean and Japanese declarative sentence-final forms: From the perspective of semantics and pragmatics. *Information*, 18-5, 1609–1622. Japan: International Information Institute.

Moon, C. (2015b). A contrastive study of Korean and Japanese inference forms, *Information,* 18-11, 4513-4525. Japan: International Information Institute.

Nihongo Kizyutu Bunpoo Kenkyuukai. (2003). *Gendai Nihongo bunpoo 4. Modariti* [Modern Japanese grammar 4. Modality]. Tokyo: Kurosio Shuppan.

Nitta, Y. (1997). *Nihongo bunpou kenkyu zyozetu* [Preliminaries to the study of Japanese grammar]. Tokyo: Kurosio Syuppan.

Umeda, H. (1990). tyosengo to nihongo no zyutugo no koozoo [A structure of predicate in Korean and Japanese]. *Nihongo Kyooiku*, *72*, 42–52.

Abbreviations

ASP   Aspect
ATT   Attributive
COP   Copula
DECL   Declarative
GOAL   Goal
INTER   Interrogative
LOC   Locative
NEG   Negative

NM   Nominative
NMLZ   Nominalizer
PAST   Past
TOP   Topic
QUOT   Quotative
NEG   Negation
SPEC   Speculation
SPEECH   Speech style

Acknowledgement

Conveying the self in a foreign language: Exploring JFL learners' self-introduction

Kiyomi FUJII
Kanazawa Institute of Technology, Japan
kfujii@neptune.kanazawa-it.ac.jp

Naomi YANAGIDA
Hitotsubashi University, Japan
n.yanagida@r.hit-u.ac.jp

Abstract

In this study, the authors researched how learners of Japanese as a foreign language introduce themselves and what they want to convey, as well as what kind of skills the speakers need to effectively communicate these points to the addressee. The self-introduction video data were collected from a project in which JFL learners interacted with EFL learners through Facebook. The survey data included learners' evaluations of their own self-introductions. Interviews were conducted with students in counterpart schools to evaluate the learners' self-introductions. The authors analyzed the data and extrapolated the most useful skills for effective conveyance from the most successful self-introductions. The results showed discrepancies between learners' self-evaluations and the opinions of the addressees. Non-verbal communication strategies were also shown to play an important role in successful conveyances. In this paper, we provide a brief overview of the project and report salient results obtained through the analysis of the data. We also share pedagogical implications of the results, and suggest alternative approaches to language pedagogy.

Keywords: learners of Japanese as a foreign language; self-introduction; self-evaluation

Povzetek

V raziskavi avtorici raziščeta, na kakšen način se učenci japonskega jezika predstavijo in kaj želijo pri tem povedati. Obenem analizirata, katere spretnosti učenci potrebujejo, da lahko uspešno podajo željene informacije. V raziskavi so bili uporabljeni videoposnetki samopredstavitev iz projekta, v katerem so učenci japonskega jezika komunicirali z učenci angleškega jezika preko platforme Facebook. Zbrani podatki vključujejo samoevalvacije učencev in intervjuje z njihovimi sogovorniki, v katerih so tudi sogovorniki podali oceno samopredstavitev. Avtorici sta analizirali podatke in iz najboljših samopredstavitev predvideli najbolj uporabne spretnosti podajanja informacij. Rezultati so pokazali precejšnjo razliko med samoevalvacijami in ocenami, ki so jih o

samopredstavitvah podali sogovorniki. Iz podatkov je bilo tudi razvidno, da neverbalna komunikacija igra pomembno vlogo pri uspešnosti podajanja informacij. Članek predstavi splošen pregled projekta in najpomembnejše rezultate analize, ki jih avtorici na koncu pedagoško ovrednotita ter na osnovi le-teh predlagata alternativne pristope k poučevanju tujega jezika.

Ključne besede: učenci japonščine kot tujega jezika; samopredstavitev; samoevalvacija

1   Introduction

Language education needs to respond to individual needs, but so far, many language lessons focus on providing learners particular sets of patterns. For example, there are sets of self-introduction patterns for which learners essentially fill in the blanks to "create" their own self-introductions. However, this activity rarely transcends a template word game and is far from a creative endeavor. Learners who want to express their own personalities are often limited by those set patterns and cannot effectively convey their own individuality. As a result, almost everyone begins to sound like the same person.

In the present study, the researchers interrogate the current language-learning environment and argue for a syllabus that reflects individual needs, in contrast to the more common functional syllabus, or even the situational (*bamen*) syllabus (Kobayashi, 2005, 2016; Sato et. al., 2015). We focus on JFL learners' self-introductions and what they try to convey, as well as how learners introduce themselves and what kind of skills the speakers need to effectively communicate these points to the addressee.

2   Literature Review

To date, much research has been conducted on how language learners' linguistic performance is judged. The research indicates that learners' linguistic performance evaluations largely depend on the evaluator. This applies to the evaluation of JFL learners' writing (Usami, Mori, & Yoshida 2009), speech performance (Ishihara, Tsurutani, & Tsukada 2011; Saito & Akiyama 2017; Tsurutani, 2010), and conversation (Choi 2008, 2013; Morimoto, Mizukami, & Yanagida 2013; Nohara 2009, 2011; Watanabe 2005a, 2005b).

For example, Tsurutani's 2010 study indicates that native Japanese college students who did not have any contact with non-Japanese judged L2 Japanese learners pronunciation based on accuracy of timing rather than pitch accent (Tsurutani 2010). A further study examined four groups of evaluators: Japanese language teachers, non-language teachers, beginner-level JFL learners, and advanced level JFL learners to see how they evaluate JFL learner speech (pitch and timing in sentences). The results

showed that, somewhat predictably, Japanese language teachers give more strict judgments than non-teachers, and advanced learners are stricter than beginner-level learners (Ishihara, Tsurutani, & Tsukada 2011).

Watanabe (2005a) studied how native speakers who are also Japanese language teachers (n = 25) or college students (n = 25) evaluate JFL learner conversations with Japanese college students to discern criteria for positive and negative evaluation. The results showed that Japanese teachers tend to evaluate negatively for sociolinguistic competence (such as failure to use honorifics), but Japanese college students, in contrast, give positive evaluations for the same lack of honorifics. Watanabe (2005a) also showed that both teachers and students evaluate grammar and pronunciation errors negatively. On the other hand, vocabulary and expressions such as *wakamonokotoba* (youth slang) and discourse and communication competence, such as non-verbal communication (using gestures and fillers), were positively evaluated. Nohara's (2011) study also found that Japanese language teachers focus on linguistic performance, but non-teachers tend to focus more on the speakers' attitude. Other studies (e.g., Choi 2008, 2013) also examined Japanese teacher and non-teacher evaluations of JFL conversations and found similar results, namely that non-verbal communication strategies, such as paralinguistic competence, are evaluated positively by both groups.

Iwasaki (2013) investigated JFL learners' (n = 5) use of hedges such as *nanka* or *chotto,* and looked at how Japanese young people, namely college students (n 60), feel about their sociability. JFL learners' discourses were recorded before and after study abroad, and the frequency of hedges and raters' evaluations were compared. The results demonstrated that, after a study abroad, the JFL learners' use of hedges increased significantly, as did the ratings by college students. Iwasaki claimed JFL learners who use hedges frequently are able 'to socially package their messages effectively' (Iwasaki, 2013, p. 263) as well as to 'project their identity as young adults' (Iwasaki, 2013, p. 263). Although Iwasaki mainly concentrated on JFL learners rather than raters, her data provides insights into the judgments of a specific rater demographic.

The studies mentioned above indicate that the evaluations of learners' linguistic performance are contingent upon the evaluator. However, there are few studies that focus on the learners' intentions. A notable exception is Siegal's (1994, 1995, 1996) study of four adult women learning Japanese in Japan. This ethnographic research focused on language use associated with the image JFL learners wanted to present. The JFL learners were aware of the different speech styles associated with expressing politeness. As one of her participants, Mary often thought she could not express certain 'subtleties' like she could in English. In cases of learners wanting to express deference or politeness, this gap seems larger for adult learners. However, interlocutors did not necessarily view pragmatic inappropriateness as a failure (Siegal, 1995, 1996). With

regard to these issues, many scholars note the recent moves in language education away from teaching methodologies that strictly focus on standardized content (Kobayashi, 2005, 2016; Sato et. al., 2015). Noda (2005), for instance, argues that it is time to revisit language education and shift focus from grammar and teaching sets of grammatical structures to developing skills suited to individual needs. To wit, rather than teaching standardized grammatical points, content should be adjusted to better fit individual needs (Noda, 2005, p. 18). Kobayashi (2005) also supports this position, contending that so-called beginner Japanese grammar does not include the learner's point of view.

To further mine this area, we focused on JFL learner intention in the present study by examining their self-introductions along with what they were attempting to convey. We also analyzed how learners introduced themselves and what kind of skills the speakers needed to effectively communicate these points to the addressee.

3   Method

3.1   Project Overview

The project comprised exchanges between JFL and EFL learners at universities in the U.S. and Japan, lasting from September 2016 to December 2016. The purpose was to encourage students to go beyond classroom learning into voluntary, self-directed learning outside the classroom, as well as using the target language in practical and personal ways. The project contained a series of tasks involving video and written exchanges via private Japanese and English Facebook group pages created by the instructors. On the Japanese page, JFL learners posted their videos in Japanese, and native Japanese college students (EFL learners) posted their comments in Japanese. Conversely, on the English page, EFL learners posted their videos in English and American college students commented in English. In these assignments, students introduced themselves, talked about their campus, and explained their favorite activities.

The participants in this study were 66 JFL and 51 EFL learners from the exchange project. The JFL learners were enrolled into third- and fifth-semester Japanese courses in the US., and their Japanese proficiency level was novice-high to intermediate-high. The EFL learners were enrolled in a Basic English course at a science and engineering university in Japan.

3.2   Materials

3.2.1   Questionnaire

We conducted surveys of JFL learner participants, with questions tailored specifically to their video self-introductions (Appendix A). The questions were the following: 1) How would you describe "yourself"? Write the key words that capture how you see yourself, 2) Do you think you were able to show "yourself" as you intended to in a Facebook video? If yes, what kind of things did you try, or how did you show who you are in the video? If no, why do you think you could not show who you are? What was difficult about it? The questionnaire was administered in English and all questionnaire data were collected and used to create a self-description sheet.

3.2.2   Self-introduction videos

Self-introduction videos were likewise collected from the Facebook-based project described above. There were three JFL classes and three EFL classes, allowing for a one-to-one class pairing. All students created the videos to introduce themselves to their counterparts at the beginning of the project. However, the due date for the EFL side was scheduled earlier than the JFL learners' date, which gave the JFL learners the chance to see their counterparts' videos and gave them some idea as to their audience for their self-introductions.

3.3   Data collection procedure

After the questionnaire data were collected, we selected JFL learners who provided detailed answers to Question 1). Among those selected JFL learners, six answered "yes" and seven answered "no" to Question 2). We created a list of self-descriptions for the 13 learners based on Question 1). We collected these 13 learners' self-introduction videos from the dedicated Facebook page and created a single 11.5-minute video of six patterns from the introductions. To avoid showing the same order of patterns, we further created differing patterns for the order of self introductions. Additionally, we gathered questionnaire data from those 13 JFL learner self-descriptions in English. Referenced by Morimoto, Mizukami, & Yanagida (2013), all description words were listed and translated into Japanese. During this process, redundant description words were eliminated and similar words were merged into one translation. For example, the words "talkative" and "verbose" were merged into the Japanese word *oshaberina*. The list consisted of 36 Japanese words and served as a document for the Japanese students' to use when they reviewed the JFL learners' self-introduction videos.

The video clips were shown to six Japanese students (three male and three female) who participated in the project (Table 1). None of them had any non-Japanese friends

or acquaintances at that point. They were asked to mark the list on the self-description sheet based on their impressions of JFL learners. Semi-structured follow-up interviews were conducted with Japanese students, in Japanese, to discern the reasoning for their evaluations. The authors translated the interview data into English.

Table 1: Japanese students' rater profiles

Rater	Age	Sex	Major (Minor)
A	20	F	Psychological informatics
B	20	F	Psychological informatics
C	20	M	Civil and environmental engineering
D	19	F	Media information
E	20	M	Media information
F	19	M	Mechanical engineering

3.4   Data collection procedure

This study focused on JFL learners (three males and three females) who claimed they could express themselves in their videos. Table 2 lists these 6 (out of 13) learners' profiles and self-introduction video information.

We compared these learners' self-descriptions with Japanese-native students' judgments, expressed in the vocabulary list that Japanese students marked while watching JFL learners' videos. We created 14 identical sheets for each Japanese student, one for testing and 13 for JFL learner videos. The authors first independently divided the data into five categories based on NEO Personality Inventory-Revised (NEO PI-R) (Costa et al 2011; Shimonaka et al. 1998) and then consulted with each other to determine the final categorization. This inventory was originally developed to assess individual personalities (Costa, McCrae, & Kay 1995). We used this inventory primarily because of its widespread use. The NEO PI-R has 60 items to provide five basic personality factors: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. For example, the self-description "outgoing" is categorized as a positive item under extraversion, while "reserved" is categorized as a negative item under extraversion. Categorized data were first rated giving one positive item 10 points, and one negative item negative 10 points, the aggregate of which gave us the point totals. Furthermore, we standardized both groups' points to compare JFL learners' self-descriptions and Japanese students' evaluations.

A native Japanese research assistant transcribed the interview data for analysis. Based on the interview data, we divided Japanese students' evaluations according to how they made their decisions while watching the JFL leaners' videos. Following Argyle (1994) and Poyatos (1992), we divided the judgments into verbal and non-verbal communication factors. Using the JFL learner data from the questionnaire

demonstrated what kinds of skills, such as verbal expressions or non-verbal cues, they used to best represent themselves in the videos.

Table 2: JFL learner profiles and self-introduction video information

Student #	Sex	Major (Minor)	Length of video	# of sentences
1	M	Math and Art	1:08	6
2	F	Linguistics	0:55	10
3	M	Anthropology (Japanese language)	1:06	10
4	M	Science engineering (Japanese literature)	0:27	6
5	F	Computer engineering	1:25	19
6	F	Business	0:35	10

4   Results and discussion

4.1   JFL learner and rater judgments

Figure 1 shows the results of JFL learners' self-descriptions against the judgment of the raters. S stands for "student," and the numbers correspond to the student number in Table 2. The solid lines indicate the JFL learners' self-description and the dotted lines indicate the rating of the Japanese evaluators. Although most JFL learners' self-descriptions differed from the rater's evaluation results, one learner's (S3) self-description was very close.

S2

S3

S4
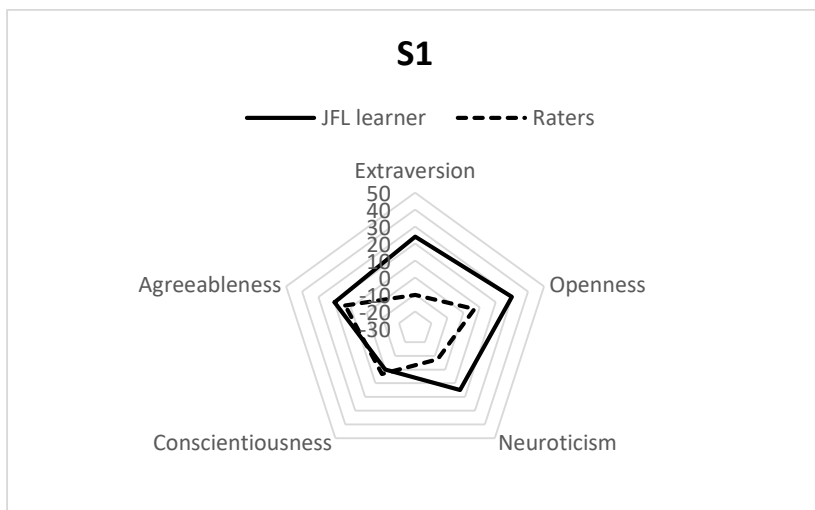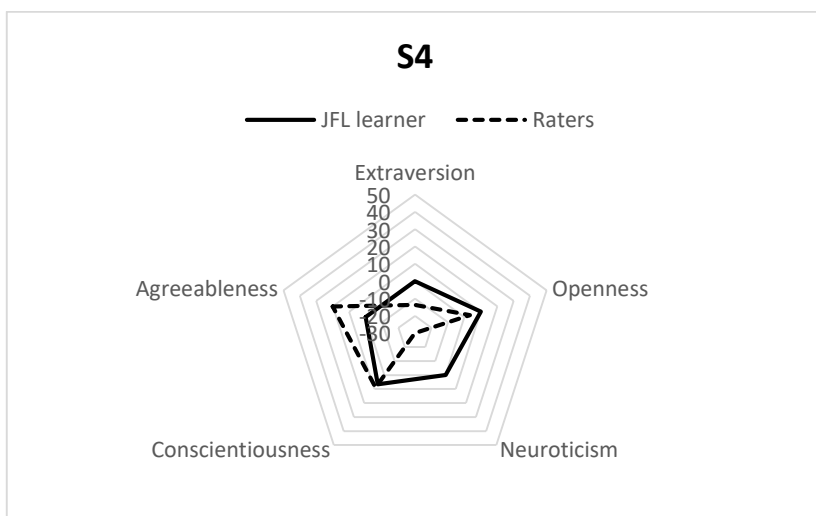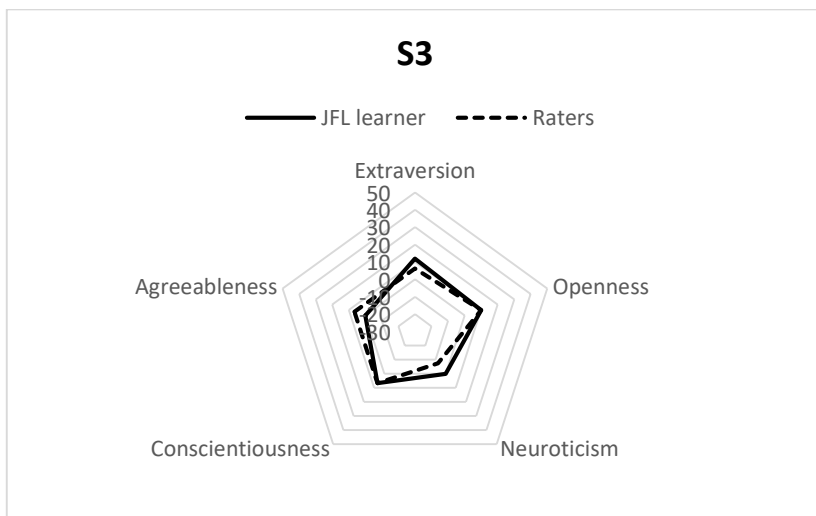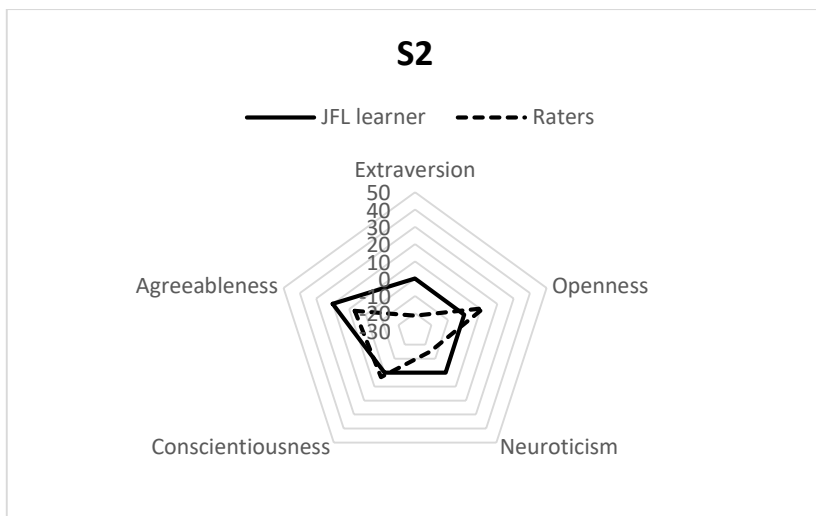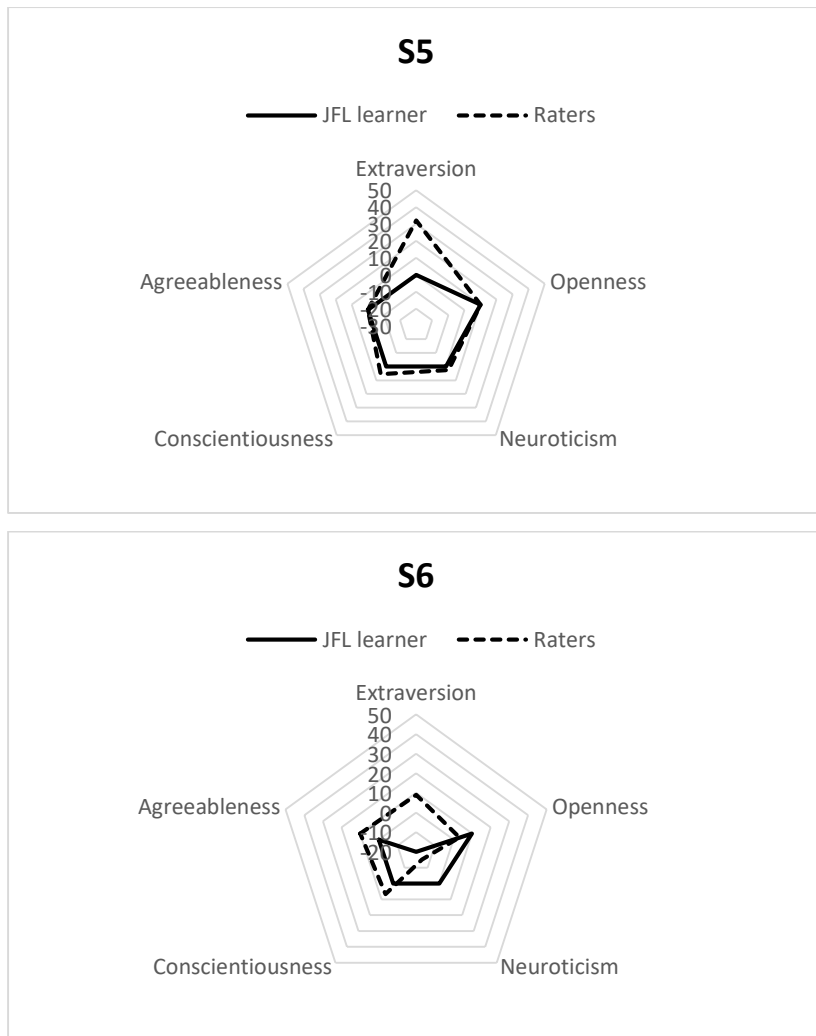
Figure 1: JFL learners' self descriptions and raters' judgments

4.2   JFL learners' intentions and evaluation factors

Based on the Japanese students' rating data, which we collected during a follow-up interview, we examined the factors behind their judgments. Figure 2 illustrates the results. Each JFL learner's self-description was examined based on whether the raters' decisions were made on verbal or non-verbal communication factors. For example, if the raters' evaluations were based on a story content, we categorized it as a verbal factor, but when the evaluations were based on paralinguistic aspects such as gestures or learner's appearance, we categorized these as non-verbal factors. As shown in Figure 2, most judgments were motivated by non-verbal factors.

The majority of comments we received from the raters had to do with paralinguistic factors, tone and volume of voice, and speed of speech, covered under 'the way they talk' (39%). At the interview we confirmed the specific rating for 'the way

they talk.' In this instance 'speaking slowly' resulted in the rating of 'calm,' whereas 'speaking energetically' led to the rating of an 'energetic person.' The raters also rated JFL learners' facial expressions (9%) and gestures (7%). For example, smiling was taken to be a sign of a 'friendly person,' and bowing at the beginning or end of the self-introduction resulted in a rating of 'respectful person.' The rater's decisions were made by mostly non-verbal factors as the previous researchers indicated.

In the next section, we focus on two cases in which learner's self-description and raters' judgments differed the most (S1) and the least (S3).
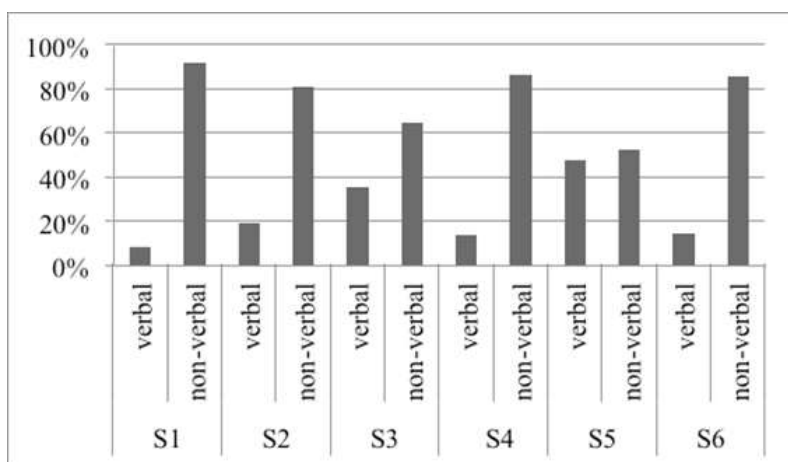
Figure 2: Factors for raters' judgments

4.3   Qualitative analysis

As we previously described, the interview data discussed in this section was translated from Japanese to English by the authors. We can see from Figures 1 and 2, in instances where evaluations were motivated by non-verbal factors, the learners' self-descriptions and raters' judgments differed greatly. For example, S1 self-described as 'loud, happy, busy, high-strung, independent, creative, energetic, smart, verbose, caring,' mentioning, 'I was able to give my answers my own personal spin. I used gesticulations similar to how I do in English. I felt fairly confident as a communicator.' (S1) However, when the raters were questioned about their evaluations, they responded: 'I thought S1 was very shy because S1 looked away (rater C).' or else 'S1 spoke slowly and did not move much, so I thought S1 was a very quiet person (rater D).' Other raters said, 'When S1 made a big gesture, I thought this was a funny person (*omoshiroi hito*) (rater E).' and 'Since talking about doing comedy, I thought S1 was a creative person (raters B & F). '. S1's method of 'using gesticulations' did not work as much as she or he expected.

On the other hand, when the raters' judgments were informed by verbal factors, the learner's self-description and the raters' judgments were much more in line, such as S3 and S5. S3's self-description was 'nice, respectful, curious, fun-loving,' as S3 explained, 'I tried to show that I was interested in differences in personal experience between myself and Japanese students. I tried to do this by asking questions about their experiences (S3).' The raters commented on non-verbal factors. Rater B for example said that 'As soon as S3 started to speak S3 was really energetic, and from the way she or he spoke, S3 was very friendly'. Some raters were also motivated by S3's verbal factors: 'S3 talked about learning dialects and I thought S3 was a very fun and friendly person (rater C)' and 'S3 is a very curious person since s/he is studying *kansai ben* (kansai dialect) (rater F).' One rater added that 'S3 called out to us saying to work hard,which made me think S3 was a hard working person (rater F).'

It is clear from the raters' comments about S1 and S3 that both non-verbal and verbal factors strongly affect their judgments. This indicates that for an effective self-introduction (here defined as matching closely with the native-speaker evaluations), it is crucial for JFL learners to develop skills that enable them to verbally express their personality rather than simply memorizing set grammatical patterns, which seems to be the standard approach. These results also suggest that JFL learners' intuitions about effective use (here defined as matching closely with the native-speaker evaluations) of non-verbal communication are generally unreliable. This seems to hold true at least at this level of Japanese ability. It would be instructive to conduct a similar study with higher-level JFL learners to see if their use of non-verbal strategies proves more successful, especially since higher-level students are also likely to have more cultural – not strictly linguistic – exposure. Nevertheless, there is a way to train such intuitions in a systematic manner. We believe that the syllabus that corresponds with individual needs (Kobayashi, 2005, 2016; Sato et. al., 2015) provides the tools that learners need to express themselves.

5   Conclusion

In this study, we investigated JFL learners' self-introductions, along with what they were attempting to convey, to shed light on how learners introduce themselves, and what kind of skills the speakers need to effectively communicate these points to the addressee. The results support findings of previous studies, showing that non-verbal factors play an important role in L2 learners' utterances. However, our results also indicated that the learners' self-description and raters' judgments were closest when judgments were based on verbal aspects. For language learners, self-introductions often constitute the first steps and may seem easy for non-novice learners, but it might be worth devoting more attention and practice to what it is one truly wants to convey. We believe a syllabus that focuses more on the individual (Kobayashi, 2005, 2016; Sato

et al., 2015) provides one of the most effective ways for language learners to achieve this goal. Although the current study suggests that learners need both non-verbal and verbal skills to effectively communicate, the data set was rather small, and we need to continue collecting data and conducting research in this area.

Acknowledgement

References

Argyle, M. (1994). *The psychology of interpersonal behavior*. Penguin UK. ISBN 0140273742

Choi, M. 崔文姫 (2008). Kankokujin nihongogakushûsha no hatsuwa nitaisuru nihongobogowasha no inshôkeisei: kotonaru zokusei wo motsu bogowasha no hyûka no sûi 韓国人日本語学習者の発話に対する日本語母語話者の印象形成: 異なる属性を持つ母語話者の評価の相違 [Japanese native speakers' assessment of Japanese learners' discourse: how Koreans learning Japanese impress Japanese people and what factors are involved]. *Nihongokenkyû* 日本語研究*, 28,* 1-15. ISSN 03866084

Choi, M. 崔文姫 (2013). Nihongogakushûsha no hatsuwa nitaisuru nihongobogowasha no hyôka: Nihongokyôshi to hinihongokyôshi no ingamoderu wo shushinni 日本語学習者の発話に対する日本語母語話者の評価:日本語教師と非日本語教師の因果ﾓﾃﾞﾙを中心に [Japanese native speakers' evaluation of Japanese learners' utterances : Cause-and-effect models for Japanese teachers and non-teachers], *Kokuritsu kokugo kenkyûjo ronshû* 国立国語研究所論集 *National Institute for Japanese Language and Linguistics Research Papers*, *5*, 1-16. ISSN 2186-134X. DOI 10.15084/00000501

Costa, P. T. Jr., McCrae, R. R., & Kay, G. G. (1995). Persons, places, and personality: Career assessment using the revised NEO personality inventory. *Journal of Career Assessment, 3 (2),* 123-139. DOI 10.1177/106907279500300202

Costa, P. T. Jr., McCrae, R. R., Shimonaka, Y., Nakazato, K., Gondo, Y., & Takayama, M. Paul T. Costa・Robert R McCrea・下仲順子・中里克治・権藤恭之・高山緑(2011). *Nihonban NEO-PI-R, NEO-FFI shiyômanyuaru* 日本版 NEO-PI-R, NEO-FFI 試用ﾏﾆｭｱﾙ *[NEO-PI-R, NEO-FFI manual for the Japanese version Big5 personality inventory]* Tokyo Japan: Tokyoshinri 東京心理. ISBN 4924552011

Ishihara, S., Tsurutani, C., & Tsukada, K. (2011). What constitutes 'good pronunciation'from L2 Japanese learners' and native speakers' perspectives? A perception study. *Electronic Journal of Foreign Language Teaching, 8 (Suppl. 1),* 277-290. ISSN 0219-9874

Iwasaki, N. (2013). Getting over the hedge: Acquisition of mitigating language in L2 Japanese. In C. Kinginger, (Ed.), *Social and Cultural Aspects of Language Learning in Study Abroad. Language Learning & Language Teaching 37*, (pp. 239-267). Philadelphia Amsterdam: John Benjamins Publishing Company. ISBN 9789027213167

Kobayashi, M. 小林ミナ (2005). Komyunikeeshon ni yakudatsu nihongokyôikubunpô コミュニケーション に役立つ日本語教育文法. In H. Noda (Ed.) 野田尚史(編), *Komyunikeeshon notameno nihongokyôikubunpô* コミュニケーションのための日本語教育文法, (pp. 23-42). Tokyo, Japan: Kuroshio Publishers くろしお出版. ISBN 4874243347

Kobayashi, M. 小林ミナ (2016). Fukugengo, fukubunkajidai no nihongokyôiku niokeru nihongokyôshiyôsei 複言語・複文化時代の日本語教育における日本語教師養成. In H. Honda & M. Matsuda 本田弘之・松田真紀子(Eds.), *Fukugengo, fukubunkajidai no nihongokyôiku: watashitachinokotobatoha?* 複言語・複文化時代の日本語教育:わたし たちのことばとは?[*Japanese language education in the era of plurilingualism and pluriculturalism*] (pp.135-162). Tokyo, Japan: Bonjinsha 凡人社. ISBN 9784893589125

Morimoto, I., Mizukami, E., & Yanagida, Y. 森本郁代・水上悦雄・柳田直美(2013). Ryûgakusei ni yoru hanashiai ni tai suru hyôka ni eikyô o ataeru komyunikeeshon kôdô 留学生によ る話し合いに対する評価に影響を与えるコミュニケーション行動 [The effect of communication behaviors on the evaluation of group discussions by Japanese language learners]. *Sôgô seisaku kenkyû* 総合政策研究[*Journal of Policy Studies*]*, 44*, 41-52. ISSN 1341996X [http://hdl.handle.net/10236/11400]

Noda, H. 野田尚史 (2005). Komyunikeeshon notameno nihongokyôikubunpô no sekkeizu コミュ ニケーションのための日本語教育文法の設計図. In H. Noda 野田尚史 (Ed.), *Komyunikeeshon notameno nihongokyôikubunpô* コミュニケーションのための日本語教育文法 (pp. 1-20). Tokyo, Japan: Kuroshio Publishers くろしお出版. ISBN 4874243347

Nohara, Y. 野原ゆかり (2009). Hatsuwa no 'wakariyasusa' o handan suru yôin: Ippan Nihonjin to bogowasha nihongokyôshi no hikaku o tooshite. 発話の「分かりやすさ」を判断する 要因―一般日本人と母語話者日本語教師の比較を通して [Factors that determine explicitness of speech: A comparison of Japanese general public and native Japanese language teachers]. *Ningen bunka sôsei kagaku ronsô* 人間文化創成科学論叢 [*Journal of the Graduate School of Humanities and Sciences*]*, 11*, 165-174. [http://hdl.handle.net/10083/34655]

Nohara, Y. 野原ゆかり (2011). Nihongogakushûshano kôtôsetumaini taisuru bogowashano hyôka: hyôkakomentono shitsutekibunsekiniyoru jireikenkyû 日本語学習者の口頭説明 に対する母語話者の評価:評価コメントの質的分析による事例研究[Native speakers' evaluations of Japanese learners' utterances: Case studies with qualitative analyses of evaluation Comments]. Hikakubunkakenkyû 比較文化研究[*Journal of the Japan Association of Comparative Culture*]*, 99*, 69-80.

Poyatos, F. (1992) The audible-visual approach to speech as basic to nonverbal communication research. In Fernando P. (Ed.). *Advances in Nonverbal Communication: Sociocultural, Clinical, Esthetic and Literary Perspectives.* (pp. 41-57). Amsterdam: John Benjamins. ISBN 9027220859.

Saito, K., & Akiyama, Y. (2017). Linguistic correlates of comprehensibility in second language Japanese speech. *Journal of Second Language Pronunciation*, 199-217. DOI 10.1075/jslp.3.2.02sai

Sato, S., Takami, T., Kamiyoshi, U., & Kumagai, Y. 佐藤慎司・高見智子・神吉宇一・熊谷由理 (2015). Miraiwotsukuru kotobano kyôikuwomezashite-naiyôjûshino hihantekigengokyôiku Critical Content-Based Instruction no riron to jissen 未来を創るこ とばの教育をめざして–内容重視の批判的言語教育 Critical Content-Based Instruction

の理論と実践 [Language education for social future: Critical Content-Based Instruction]. Nihongokyôikugakukenkyû 日本語教育学研究 6. Tokyo, Japan: Koko shuppan ココ出版 [Coco Publishing].

Shimonaka, Y., Nakazato, K., Gondo, Y., & Takayama, M. 下仲順子・中里克治権藤恭之・高山緑 (1998) NihonbanNEO-PI-R no sakuseito sono inshitekidatouseino kentô 日本版 NEO-PI-R の作成とその因子的妥当性の検討[Construction and factorial validity of the Japanese NEO-PI-R]. *Seikakushinrigakukenkyû* 性格進学研究[*The Japanese Journal of Personality*]*, 6 (2),* 138-147.

Siegal, M. (1994). Second-language learning, identity, and resistance: White women studying Japanese in Japan. *Culture Performances*, *Proceedings of the Third Berkeley Women and Language Conference*, (pp. 642-650). Berkeley, CA: Berkeley Women and Language Group. ISBN 10 999773596X

Siegal, M. (1995). Individual differences and study abroad: Women learning Japanese in Japan. In Freed F., Barbara, (Ed.), *Second Language Aquisition in a Study Abroad Context*, (pp. 225-244), Philadelphia, PA: John Benjamins, B.V. ISBN 9781556195426

Siegal, M. (1996). The role of learner subjectivity in second language socio-linguistic competency: Western women learning Japanese. *Applied Linguistics, 17*, 356-382. DOI https://doi.org/10.1093/applin/17.3.356

Tsurutani, C. (2010) Foreign accent matters when timing is wrong. In Proceedings of Interspeech 2010, (pp.1854–1857). http://hdl.handle.net/10072/36647, Accessed 5 Feb 2019

Usami, Y., Mori, A., & Yoshida, S. 宇佐美洋・森篤嗣・吉田さち (2009). Gaikokujin ga kaita nihongotegamibun ni taisuru nihonjin no hyôkataido no tayôsei; shitsutekishuhô niyoru keesusutadei. 外国人が書いた日本語手紙文に対する日本人の評価態度の多様性: 質的手法によるケーススタディ[Diversity in Japanese native speaker attitudes when evacuating letters written by non-natives: a qualitative method case study]. *Shakaigengokagaku* 社会言語科学 [*The Japanese Journal of Language in Society*]*, 12 (1),* 122-134. ISSN 13443909

Watanabe, T. 渡部倫子 (2005a). Purasuhyôka, mainasuhyôka sareyasui hatsuwa no yôso towa: Nihongogakushûsha nitaisuru nihongowasha no hyôka. プラス評価・マイナス評価されやすい発話の要素とは:日本語学習者に対する日本語母語話者の評価 *Kyôikugakukenkyûjânaru* 教育学研究ジャーナル[*The Journal of Educational Research*]*, 1,* 77-81. DOI 10.20677/csssej.1.0_77

Watanabe, T. 渡部倫子(2005b). Nihongogakushûsha no hatsuwa nitaisuru nihongobogowasha no hyôka: kyôbunsankôzôbunseki niyoru hyôkakijun no kaimei 日本語学習者の発話に対する日本語母語話者の評価―共分散構造分析による評価基準の解明. *Nihongokyôiku* 日本語教育 [*Journal of Japanese Language Teaching*]*, 125*, 67-75. ISSN 03894037

Yoshida, S. 吉田さち (2014). Nihongogakushûsha no gengo un'yô nitaisuru nihongobogowasha no hyôka: Bamen niyori bogowasha no hyôka wa kotonaruka 日本語学習者の言語運用に対する日本語母語話者の評価: 場面により母語話者の評価は異なるか [Native Japanese speakers' evaluations of Japanese-learners' linguistic performance: Do Native Japanese speakers' evaluations differ according to the

situation?]. *Komyunikêshonbunka* コミュニケーション文化[*Communication in Culture*]*, 8*, 27-43. ISSN 18818374

Appendix A

Facebook Project Reflection Log

About Video #1: Self-Introduction-Converying the self in a foreign language

1. How would you describe yourself? Write the key words that capture how you see yourself.

2. Do you think you were able to show "yourself" as you intended to in a Facebook video?
   Yes / No

3. If YES → What kind of things did you try, or how did you show who you are in the video? (verbal/non-verbal/others)
   If NO → Why do you think you couldn't show who you were? What was difficult about it?

Retrieving Linguistic Information from a Corpus on the Example of Negation in Chinese

Ľuboš GAJDOŠ

Comenius University in Bratislava, Slovakia

lubos.gajdos@uniba.sk

Abstract

The paper deals with corpus analysis of negation in Chinese, namely the negatives bù 不 and méi/méiyǒu 没/没有. The adverbs BU and MEI are two of the most frequent negatives in Chinese. The aim of this study is to present statistical data together with linguistic analysis. The results provide empirical evidence of discrepancy between "authentic" language data versus linguistic prescription with practical implications for second-language acquisition. The findings inter alia suggest a new approach to verb categorisation.

Keywords: Chinese language; corpus linguistics; quantitative description; negation; potential complements

Povzetek

Članek obravnava korpusno analizo negacije v kitajščini, pri čemer se avtor osredotoča na prislova bù 不 in méi/méiyǒu 没/没有, ki sta najpogostejši nikalnici v sodobnem kitajskem jeziku. Namen prispevka je predstaviti statistične podatke v povezavi z jezikoslovno analizo. Rezultati študije prinašajo empirične dokaze o neskladju med jezikovno rabo in jezikovnimi normami, ta spoznanja pa je moč uporabiti tudi pri poučevanju kitajščine kot tujega jezika in za premislek o drugačnem pristopu na kategorizacijo glagolov.

Ključne besede: kitajščina; korpusno jezikoslovje; kvantitativni opis; negacija; zmožnostna dopolnila

1   Introduction

Generally speaking, there are a number of negatives in modern Chinese.[1] In this article only two negative adverbs, namely *bù* 不 and *méi/ méiyǒu* 没/没有[2] (hereafter referred to as BU and MEI), are discussed. The Hanku corpus is used[3] as the primary source of language material and statistical data. As the intention is to mainly use the corpus-driven[4] approach to studying of negation, thus the previous linguistics research on this topic is left aside.

Let us start with some basic queries:[5]

[tag="AD" & word="不"]
[tag="VV|AD" & word="没|没有"][6]

The results are 7371142 (9897.85 per million),[7] 686352 (921.62 per million)[8] respectively. These numbers only tell that the token BU is approximately 10 times more frequent than MEI. The difference is even more pronounced when searching in a certain variety of Chinese, e.g. in the corpus of legal Chinese, the occurrence of BU is 45254 (6 281.56 per million), the occurrence of MEI 720 (99.94 per million). Let us take a closer look at the tokens that collocate with these negatives. The following queries should return collocates at the position 1 on the right side:[9]

[tag="AD" & word="不"][]
[tag="AD"& word="没|没有"][]

---

[1] For details see e.g. Liu (2004, pp. 253–258).

[2] For the sake of simplicity, both negatives *mei* and *meiyou* are treated as two forms of one negative, namely MEI. On the other hand, their collocative partners may differ because of e.g. prosodic factors.

[3] See more in Gajdoš, Garabík and Benická (2016, pp. 53–65).

[4] See more in Baker, Hardie and McEnery (2006, p. 49).

[5] In this article, the Corpus Query Language (hereafter CQL) is used to search for collocations. With CQL, complex criteria can be set to find one or many tokens. Criteria for each token must be between a pair of square brackets [ ], e.g. [attribute="value"]. See more at https://www.sketchengine.eu/documentation/cql-basics/

[6] As there are more tags (e.g. VV = verbs, VE = YOU as the main verb, AD = adverbs) dedicated to tokens *mei* 没 and *meiyou* 没有, it is rather difficult to accurately determine the value of the negative MEI. Thus, I only use tags VV, AD in this article. For more details on the tagset see Fei (2000, pp. 4–35).

[7] Unless stated otherwise, frequencies are given in absolute occurrence in the Hanku corpus.

[8] The occurrence of *mei* 没 is 342190 (459.49 per million) and 344162 (462.14 per million) for *meiyou* 没有.

[9] The regular expressions may match the following patterns:

adverb *bu* + any token or
adverb *mei/meyou* + any token.

As it is rather difficult to identify the collocates at the position further to the right using only POS tags, this topic will be left for future research. See e.g. Gajdoš (2018).

The results are summarised in the tables below.[10]

Table 1: The most frequent POS at the position 1 (Corpus: web-zh)

Query: word, [tag="AD" & word="不"][]
Query: word, [tag="AD|VV" & word="没|没有"][]

| 不 | | 没\|没有 | |
| No. of results: 7371142 | | No. of results: 686352 | |
tag	frequency	tag	frequency
AD VV	5092363	**VV VV**	326984
AD VA	827140	**AD VV**	193781
AD VC	695511	**VV AD**	35762
AD AD	444055	**VV P**	30138
AD P	155025	**AD AD**	16246
AD PU	25581	**VV AS**	15136
AD AS	19730	**VV PU**	9902
AD JJ	19525	VV CD	5771
AD PN	13121	VV PN	5350
AD NN	10044	VV VA	4975
AD BA	9276	VV BA	4870
AD VE	8798	VV NN	4449
AD LB	8553	VV DT	4223
AD CD	7535	AD VA	3511
AD SB	6815	AD P	3488
AD NR	5930	VV SB	2795
AD DT	4226	VV DEC	2660
AD DEC	3170	VV LB	1914
AD DEV	2380	AD SB	1541
AD SP	1987	VV DER	1460
AD LC	1744	AD BA	1334
AD M	1709	VV JJ	1209
AD MSP	1545	VV NR	1106
AD NT	1097	VV NT	959
AD OD	1032	AD CD	861
AD CC	1023	AD PN	691
AD CS	1021	AD VC	663
AD DEG	941	VV VC	526
AD DER	118	VV OD	499
AD ETC	88	VV LC	380
AD FW	45	AD VE	354
AD IJ	14	AD LB	338

---

[10] The results are calculated using the NoSketch Engine UI – Node tags.

The table indicates different collocability for the negative BU and MEI, e.g. the negative BU exhibits a strong preference for copulas (here VC).[11] For practical reasons, only the POS tags, which are more frequent than 1% of each group (here in bold), are included in the analysis. The PU tag is also to be excluded from further analysis as it stands for punctuation.

Table 2 shows 10 of the most frequent collocates for each negative. The results are calculated using the NoSketch Engine UI – Node forms.[12]

Table 2: The most frequent tokens at the position 1 (# Corpus: web-zh)

Query: word, (meet [tag="VV|VA|VC|AD|P"]2:[tag="AD" & word="不"]-1 -1)
Query: word, (meet [tag="VV|AD|P|AS"]2:[tag="AD|VV" & word="没|没有"]-1 -1)

| 不 | | 没\|没有 | |
| No. of results: 7214094 | | No. of results: 621555 | |
word	frequency	word	frequency
是	693223	想到	56260
能	377529	看	14539
会	311133	**能**	13999
知道	288756	了	12223
要	227692	见	9843
存在	125328	在	9249
可	125212	说	9215
到	124252	想	9141
得	91523	看到	7189
敢	85127	用	6744

At first sight, it is surprising that the collocation MEI+ *néng* 能 is the third most frequent, despite the fact that most grammars and textbooks deny this possibility.[13] Similar findings may provide the impetus for further research which would take greater

---

[11] The co-occurrence of MEI+VC is caused by the misspelling of the character *shi* in most cases, e.g. *mei shi* 没是 instead of *mei shi* 没事.

[12] To find collocative partners of both negatives, the operator *meet* is used. That means that the corpus is search for the following patterns:

   adverb BU + verbs (VV) or
   adverb BU + adjectives (VA) or
   adverb BU + copulas (VC) or
   adverb BU + adverbs (AD) or
   adverb BU + prepositions (P).

See more at https://www.sketchengine.eu/documentation/cql-meet-union/

[13] There are some exceptions, e.g. Švarný and Uher (2014, p. 48) describe this phenomenon and Liu (2004, p. 257) also suggests this possibility, however, they do not further elaborate this point.

account of actual language use. The "new" grammars or textbooks should be then based on such research.

After searching the first hundred examples manually, it turns out that the co-occurrence of some tokens with BU is higher than one would expect based on the frequency of affirmation, e.g. *liǎo* 了 (70691), *zhù* 住 (65400), *qǐ* 起 (40086) etc., furthermore, these verbs typically serve as so-called complements.[14] That means that only tentative conclusion may be drawn from this evidence, nevertheless, it should play a role when comparing the overall frequency of both negatives. I discuss this topic further in the chapter *Potential complements*.

2   Potential complements

Let us return to the examples that have been mentioned in Chapter 1 and analyse them.

(1) 他   办   不**/AD**   了**/VV**   此事
    Tā   bàn   bù/AD   liǎo/VV   cǐ shì
    He   manage   BU-neg.   compl.   this matter
    'He cannot do this.'

(2) 杰克终于   忍   不**/AD**   住**/VV**   说了
    jiékè zhōngyú rěn   bù/AD   zhù/VV   shuō le
    Jack finally   endure   BU-neg.   compl.   speak LE
    'Jack finally couldn't help saying it.'

(3) 这辈子房子   买   不**/AD**   起**/VV**   了
    zhè bèizi fángzi mǎi   bù/AD   qǐ/VV   le
    this life house   buy   BU-neg.   compl.   LE
    '(One) cannot afford to buy a house for the entire life.'

(4) 对于双方不   能   不**/AD**   说**/VV**
    duìyú shuāngfāng bù néng   bù/AD   shuō/ VV
    for to both sides not able to BU-neg.   speak
    'regarding (things that) both sides cannot but speak'

Randomly selected samples suggest that many examples may be considered as so-called potential complements with the "morphological" structure VV + BU + VV while

---

[14] See e.g. Yip (2009, pp. 234–241).

the first morpheme (verb) is not equal to the third. The following query meets this condition:[15]

(meet (meet 1:[tag="VV"][tag="AD" & word="不"]-1 -1) 2:[tag="VV|VA"]-2 -2) & 1.word!=2.word

The examples below show that the regular expression does not always match the desired pattern and therefore must be modified.

(5) 那个人呢叫李一，　　知　　不　　　知道**/VV**？
Nàgè rén ne, jiào Lǐ Yī, zhī bù zhīdào
That person is Li Yi　　know BU-neg. know
'Do you know that that person is called Li Yi?'

(6) 您　能　　不　　　能够**/VV**　再具体地跟我们讲一下？
Nín néng bù nénggòu zài jùtǐ de gēn wǒmen jiǎng yīxià?
You able to BU-neg. able to　　tell us more specifically
'Can you tell us this again more specifically?'

(7) 一定程度上，不　　　能　不　　说**/VV**
Yīdìng chéngdù shàng, bù néng bù shuō/VV
to a certain extent, not　　able BU-neg. speak
'To a certain extent, one cannot but speak.'

(8) 一个　说　不　　要**/VV**
Yīgè shuō bù yào
one　say　BU-neg. want
'One says *no*.'

(9) 这些人　　找　不　　到**/VV**　工作
Zhèxiē rén zhǎo bù dào/VV gōngzuò
these people find BU-neg. compl. work
'These people cannot find work.'

There is a point worth noting here as well – auxiliary verbs (e.g. modal verbs) must be removed from the search pattern. As there is no dedicated tag for modal or auxiliary

---

[15] This regular expression matches the following pattern:
verb2/adjective2 + BU + **verb1** and verb1 ≠ verb2, the verb1 is KWIC (Key Word in Context).

verbs (except VE, VC), each of the verbs must be enumerated in the query with the attribute "word".[16] A double negative must be excluded too. The refined query is:[17]

> (meet (meet (meet 1:[tag="VV"& word!= "要|能" "& word="(?i).{1,2}"]
> [tag="AD" & word="不"]-1 -1) 2:[tag="VV|VA"& word!= "要|能"]-2 -
> 2)[word!="不"]-3 -3) & 1.word!=2.word

The following table shows the result. The overall frequency is 828224 (1112.13 per million).

Table 3: The most frequent potential complements – the negative form (# Corpus: web-zh)

Query: word,(meet (meet (meet 1:[tag="VV"& word!= "会|能|应该|该|必须|可以|可|应当|可以|应|能|能够|必须|须|要|可能|会|需要|愿意|敢|该|需|知道" & word="(?i).{1,2}"][tag="AD" & word="不"]-1 -1) 2:[tag="VV|VA"& word!= "会|能|应该|该|必须|可以|可|应当|可以|应|能|能够|必须|须|要|可能|会|需要|愿意|敢|该|需|知道"]-2 -2)[word!="不"]-3 -3) & 1.word!=2.word

word	Frequency
到	111457
了 *liǎo*	65697
住	59793
起	36774
得 *dé*	33857
出	31377
上	29568
开	22326
过	19412
见	12632
着 *zháo*	10366
出来	10114
懂	9877

---

[16] The query above contains only two of these verbs, the others are present here, e.g. 能|应该|该|必须|可以|可|应当|应|能够|必须|须|要|可能|会|可|需要|愿意|敢|该|需 etc. The limit for the length of the tokens is set to 1 or 2 by the expression: "word="(?i).{1,2}".

[17] The regular expression means that the corpus is searched for the following pattern:
token (not BU) + verb2 (not 要 nor 能) + adverb BU + mono- or disyllabic **verb1** (which is not 要 nor 能) and verb1 ≠ verb2. Only the verb1 is KWIC in the concordance and other tokens are used as contextual filters. See more at https://www.sketchengine.eu/documentation/cql-meet-union/

The result of the affirmative form might be achieved by the same query with only minor modification:[18]

> (meet (meet (meet 1:[tag="VV"& word!= "要|能" "& word="(?i).{1,2}"]
> [tag="DER"]-1 -1) 2:[tag="VV|VA"& word!= "要|能"]-2 -2)[word!="不"]-3 -3)
> & 1.word!=2.word

The total frequency of 167822 (225.35 per million) clearly shows that the occurrence of the affirmative form is far less frequent. This fact only validates the previous assumption mentioned in the literature.[19] The following list contains a sample of the most frequent verbs: 上 11598, 起 10084, 到 9769, 住 7 614, 出 7607, 出来, 3736, 见 2977 etc.

If we move back to the calculation of the overall frequency of BU, the value of the negative form of potential complements (1112.13 per million) should be subtracted from the total frequency, i.e. 8785.72 per million. Needless to say, these are only approximate numbers and further research is required.

3   Verb collocates

The first chapter discusses the collocability of the negative BU and MEI. In this chapter, I further explore this topic. When comparing the total frequency of BU vs. MEI, some considerations should be taken into account, i.e. some verbs/adjectives collocate with BU only, some registers use only a limited number of MEI etc.

After saving the results as a text file (from the NoSketchEngine UI), I proceed to test the 2 lists[20] for the duplication[21] and calculate the average value of co-occurrence. When comparing two lists for duplication in the spreadsheet program, there are many tokens in the MEI list which are marked as they have no counterpart in the BU list. This might cause surprise at first since one would expect only tokens from the BU list not having a counterpart. The explanation is rather simple: (1) most of these tokens have a disyllabic morphological structure (V+X), e.g. 找到, 看到 and cannot be paired with their monosyllabic counterpart in the BU list by the spreadsheet program (e.g. 找, 看) or (2) the frequency of the BU counterpart is below the lowest frequency of samples (see footnote 13).

---

[18] There is a dedicated tag for the 得 *de*-marker, i.e. DER.

[19] See e.g. Liu (2004, p. 583).

[20] Each list contains the 1000 most frequent verbs that collocate with BU and MEI.

[21] This might be done in MS Excel, LibreOffice Calc or any spreadsheet program.

Table 4: The 10 most frequent verbs collocating with BU and MEI (# Corpus: web-zh)

Query: word,(meet [tag="VV"]2:[tag="AD" & word="不"]-1 -1)
Query: word,(meet 1:[tag="VV" & word="(?i).{1,2}"][tag="AD|VV" & word="没|没有"]-1 -1)

不		没|没有	
word	Frequency	word	Frequency
能	377442	想到	56260
会	311129	看	14539
知道	288756	能	13999
要	227691	见	9843
存在	125328	说	9215
可	124875	想	9141
到	122930	看到	7189
得	91458	用	6086
敢	85108	去	5602
知	78854	来	5371

The results indicate that:

- From the list of the 1000 most frequent tokens (verbs) with the negative BU, 619 tokens collocate with MEI too, yet from the 100 most frequent tokens, there are 69 of them; the rest are e.g. the following tokens: 知, 行, 愿, 愿意, 肯, 应, 信 etc. that co-occur with BU only;

- From the list of the 100 most frequent tokens (verbs) with the negative MEI, a few preferably collocate with MEI, e.g. 发现, 料到, 必要, 开始, etc.;[22]

- The lower the frequency of a token in the BU list, the less frequent it collocates with both negatives;

- Generally speaking, the co-occurrence of the negative MEI with the same verb is about 2.5-time less frequent as with the BU negative, however, statistical data reveals great disparities between tokens (see table 5). That is to say that verbs on the left side of the table collocate almost always with the negative BU, on the other hand, verbs on the right side almost exclusively collocate with the negative MEI.

---

[22] This may be seen from the following comparison: the query [word="没|没有" & tag="VV|AD"][word="发现"] with the frequency of 4542 (6.10 per million) and the query [word="不" & tag="AD"][word="发现"] 62 (0.08 per million).

Table 5: Collocability of verbs (# Corpus: web-zh)

Preference for BU		Preference for MEI	
word	**ratio**	**word**	**ratio**
知道	1511,8	想到	0,005
存在	858,4	放松	0,182
会	781,7	看到	0,201
住	408,7	留下	0,277
可	325,2	出现	0,290
在	307,7	进入	0,314
起	301,4	选择	0,323
合	278,5	感觉	0,340
应该	236,3	受到	0,342
了	235,4	表现	0,385

4   Adjective and adverbs collocates

This chapter focuses on the collocability of adjectives and adverbs and the same searching methods are used.

As for the adjectives, a brief look at the given statistical data (827140 or 1110.67 per million vs. 8486 or 11.39 per million; see table 6) demonstrates that adjectives (almost) always collocate with the negative BU. The exceptions here may be considered as phrases.

Table 6: Collocability of adjectives (# Corpus: web-zh)

# Query: word,(meet [tag="VA"]2:[tag="AD" & word="不"]-1 -1)			
# Query: word,(meet [tag="VA"]2:[tag="AD|VV" & word="没|没有"]-1 -1)			
不		没|没有	
No. of results: 827140		No. of results: 8486	
word	Frequency	word	Frequency
好	71999	错	1978
美	58946	成功	884
多	46448	好气	812
够	34917	多	374
一样	28451	真正	329
大	23685	必要	257
美观	18118	好	251
高	14844	明确	205
容易	14831	成熟	195
小	13191	好好	180

The situation with regard to adverbs is a little different. While the results indicate a strong tendency to the negative BU, yet both negatives may be used.

Table 7: Collocability of adverbs (# Corpus: web-zh)

# Query: word,(meet [tag="AD"]2:[tag="AD" & word="不"]-1 -1)			
# Query: word,(meet [tag="AD"]2:[tag="AD|VV" & word="没|没有"]-1 -1)			
不		没|没有	
No. of results: 444055		No. of results: 52008	
word	frequency	word	frequency
再	76724	那么	6480
太	39336	再	4637
一定	22521	这么	4062
就	19298	完全	2989
只	17114	多	2855
曾	13088	怎么	2316
单	12286	真正	1889
正	10410	不	1541
多	8623	太	988
怎么	8621	甚么	920

5   Conclusion

To begin with, statistical data given in this study should only be taken as exhibiting a general tendency and not as a fully accurate description of "real" language. It should also be pointed out that this paper only examines the occurrence of negatives at the first position to the left of collocates. In this respect, new methods should be devised for solving issues addressed here, e.g. the problem with the POS annotation and its error rate which may significantly affect statistical data or the problem with identifying the difference between the negative MEI and the verb *yǒu* 有 (with the tag VE) etc. This leads us to the questions how to interpret the results in light of these points and what valuable results this study brings.

Firstly, when comparing results of both negatives, it seems that some verbs described as "auxiliary" or "modal" tend to collocate with the negative MEI more often than stated by language prescription. On the other hand, empirical data support the claim that adjectives only collocate with the negative BU. As for adverbs, there is still a strong preference for BU, but because I do not consider adverbs as a "true" collocate to negatives (rather as part of a bigger structure), this question should be explored in future research.

Let us now move on to the negative MEI. There are many verbs that preferably collocate with MEI rather than with BU. A closer look at the results reveals that their

morphological structure is disyllabic and the left morpheme is often a so-called "resultative complement" (*jiéguǒ bǔyǔ* 结果补语). This finding may imply that the category of verbal aspect and tense[23] deserves closer attention. That means if MEI is regarded as past time marker, these verbs are commonly used in past tense and the present tense (with BU) may describe the situation as a condition or future tense. A similar phenomenon is also observed in some Slavic languages, where the present and preterite of perfective verbs fulfil these functions too (e.g. compare the present perfective form "urobím" vs. the past perfective form "urobil" in Slovak). This suggests that these verbs in Chinese might be treated as *perfective*. In order to fully explore this topic, the marker *le* 了, as a counterpart to the negative MEI, should be included in an comparative analysis. There is a very detailed, corpus-based study conducted on this subject by Petrovčič (2009), *Operator Le in Chinese* worth noting here.

To conclude, the article shows how to use a corpus when searching for evidence of some language phenomena. As for negation in Chinese, the paper only suggests a different approach to this subject and additional research is needed.

References

Baker P., Hardie A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Fei, X. (2000). *The Part-Of-Speech Tagging Guidelines for the Penn Chinese Treebank (3.0)*. Available at https://www.cs.brandeis.edu/~clp/ctb/posguide.3rd.ch.pdf

Gajdoš, Ľ., Garabík, R., & Benická, J. (2016). The New Chinese Webcorpus Hanku – Origin, Parameters, Usage. *Studia Orientalia Slovaca, 15*(1), 53–65.

Gajdoš, Ľ. (2018). Korpusová analýza adverbiále v právnej čínštine [Corpus Analysis of adverbial complements in legal Chinese]. In T. Guldanová (Ed.), *Kontexty súdneho prekladu VII* (pp. 27–39). Bratislava: Univerzita Komenského.

Liu, Y. [刘月华] et al. (2004). *Practical Chinese Grammar* [实用汉语语法]. Beijing: Shangwu yishuguan.

Petrovčič, M. (2009). *Operator Le in Chinese.* Saarbrücken: VDM Verlag.

Petrovčič, M. (2017). Traditional and Contemporary Approaches to Chinese Particles. In S. Bračič & M. Petrovčič (Eds.), *Partikeln überall: Deutsch - Slowenisch - Chinesisch*, (pp. 103-122). Ljubljana: Znanstvena založba Filozofske fakultete.

Sketch Engine. Available at https://www.sketchengine.eu

Švarný, O., & Uher, D. (2014). *Prozodická gramatika čínštiny* [Prosodic Grammar of Chinese Language]. Olomouc: Univerzita Palackého.

Yip, P., & Rimmington, D. (2009). *Basic Chinese.* Abingdon: Routledge.

---

[23] See also Petrovčič (2017, pp. 108–109).

Appendix: The Hanku tagset[24]

Tag	English	Example
AD	adverb	也
AS	aspect particle	着
BA	preposition BA in ba-construction	把
CC	coordinating conjunction	和
CD	cardinal number	十五
CS	subordinating conjunction	如果
DEC	markers – nominalizer	吃的
DEG	genitive marker	他的
DER	resultative DE 得	说得
DEV	manner DE 地	公正地
DT	determiner	这
ETC	et cetera	等
FW	foreign word	ISBN
IJ	interjection	喂
JJ	other noun-modifier	女
LB	preposition BEI in long bei-construction	被
LC	localizer	上
M	measure word	个
MSP	other particle	所
NN	noun	记者
NR	proper noun	英语
NT	temporal noun	今年
OD	ordinal number	第三
ON	onomatopoeia	哈哈
P	preposition	从
PN	pronoun	我
PU	punctuation	。
SB	preposition BEI in short bei-construction	被
SP	sentence-final particle	了
VA	predicative adjective	大
VC	copula	是
VE	verb 有/没有/无 as the main verb	有
VV	verb	说

---

[24] For details see Fei (2000, pp. 4–35).

Korean Orthography of Loanwords and spelling problems with proper nouns from Slovenia

Minyoung CHANG
University of Ljubljana
oddmonica@gmail.com

Chikako SHIGEMORI BUČAR
University of Ljubljana
Chikako.bucar@guest.arnes.si

Abstract

This paper firstly introduces The Korean Orthography of Loanwords and its history. Recently with more cultural and economic exchanges between Korea and Slovenia, the Korean Orthography guidelines for Slovene words are needed to guide words of Slovene origin when settling in Korean. Although Serbo-Croatian is not an official language of Slovenia, some Slovene words are currently written in Korean according to the guidelines for Serbo-Croatian in KOL. Therefore, we exam the current Korean orthography examples of Slovene words and suggest amendments based on the origin pronunciation. In order to avoid further confusion and to provide the foundation of Korean Orthography of Slovene words, we firstly demonstrated a comparison table of Slovene alphabet (abeceda) and Korean alphabet (Hangeul) supported with examples of Slovene words, and then suggested detailed guidelines on how to write Slovene words in Hangeul.

Keywords: Korean orthography; Korean; Slovene; loanwords; foreignwords; spelling problems with Slovene words in Hangeul

Povzetek

Članek najprej predstavi Korejski zapis prevzetih besed and zgodovino prevzemanja. S stopnjevanjem kulturnih in ekonomskih izmenjav med Korejo in Slovenijo se je povečala tudi potreba po pravopisnih pravilih, ki bi besede slovenskega izvora primerno uvrstila v korejski zapis. Srbohrvaščina sicer ni uradni jezik v Sloveniji, vendar so trenutno korejske besede, ki so izvorno slovenske, zapisane po pravopisnih pravilih, ki so postavljena za izvorno srbohrvaške besede. V raziskavi torej najprej analiziramo primere zapisa slovenskih besed v korejščini in predlagamo spremembe, s katerimi bi korejski zapis bolj približali izgovarjavi izvorne besede. Z namenom, da bi se izognili nadaljnji zmedi ter da bi podali osnovo pravopisnim pravilom za slovenske besede v korejščini, smo najprej primerjali slovensko abecedo in korejski zapis *hangul* in dodali primere, zatem pa predlagali natančna pravila za zapis slovenskih besed v korejščini.

Ključne besede: korejski zapis; korejščina; slovenščina; prevzete besede; tujke; problemi pri zapisovanju slovenskih besed v hangulu

1   Introduction

Recently, interest for Slovenia in Korea is increasing as there are more cultural and economic exchanges between the two countries. Slovenia has been exposed to Koreans by mass media more and more. There were documentaries and traveling TV shows introducing beautiful, peaceful and green Slovenia to Koreans. Moreover, Slovenia became production locations for Korean television series [1] and for a real variety show [2]. In addition, SNG Drama Ljubljana (Slovene National Theatre) performed plays in Korea, *The Crazy Locomotive (폭주 기관차)* in 2015 and *Faust (파우스트)* in 2016 as they were invited to the Seoul Performing Arts Festival for two consecutive years. Conversely, Korean animated children's television series *Robocar Poli (로보카 폴리)* has been airing for more than a year in Slovenia. In the field of publication, several books on Slovenia and Ljubljana have been published recently in Korea. [3] Interexchange is also very active in the filed of sports, where during the PyeongChang 2018 Winter Olympic Games many Slovene athletes played well and gained the attention of Koreans. Mitja Gasparini [4], a Slovene volleyball player for example, is very popular in the V-League (South Korean club volleyball competition). Getting more and more familiar to Koreans, Slovenia became a fascinating tourist destination, with the number of tourists from Korea increasing every year.

Korean Studies of the Department of Asian Studies at the Faculty of Art, University of Ljubljana, have been receiving questiones on how to pronounce and write Slovene words in Korean by institutions and public organizations as well as individuals. This suggests that Slovene words are spoken and written among Koreans for their communication not only among people of certain backgrounds but in general public. Being aware of the need, the Orthography Commission [5] The Slovenian Academy of Sciences and Art [6] in cooperation with the Department of Asian Studies at the Faculty of Arts, University of Ljubljana, proposed the guidelines on how to write and pronounce Korean names in Slovene based on Korean Romanization. [7] The guidelines were published in February 2018 on the occasion of PyeongChang 2018 Winter Olympic

---

[1] Filming of *Dear My Friends (디어 마이 프렌즈)* in 2016, which had more than 5% audience ratings at average and *Black Knight: The Man Who Guards Me (흑기사)* in 2017, which had more than 10% audience ratings at average.

[2] Filming of *Hiring a village for a short time (잠시만 빌리지)* was in 2018. A real variety show is a combination of reality and variety program. This real variety show showed how a married couple with their baby lived in Slovenia for 2 weeks. The married couple were Korean singers, and they sang and played guitar in Vlado Kreslin's concert.

[3] Kang Byoung Yoong. (2015). Ljubljana, a city that resembles my wife (류블랴나, 아내를 닮은 도시). Paju: Nanda. Kim Yideum. (2016). *Dear Slovenia (디어 슬로베니아).* Goyang: Logopolis.

[4] Mitja Gasparini belongs to Incheon Korean Air Jumbos.

[5] Pravopisna komisija, 맞춤법 위원회

[6] The Slovenska Academija Znanosti in Umetnosti, , 슬로베니아 학술원

[7] Cf. Provapisna Komisija Pri SAZU in ZRC SAZU http://pravopisna-komisija.sazu.si/Priporočila.aspx

Games. What is also of great importance and still needs to be done is to set Korean orthography guidelines for words of Slovene origin.

Korea and Slovenia carry a common view that their languages should preserve some distance towards the flood of foreign words coming mostly from English, as young generations consume foreign words without repulsion. Some groups of people fear that with such tendencies language destruction is unavoidable. In this regard it is meaningful that Korean orthography of loanwords gets attention in Sloveniad, and that Slovenes get familiar with principles on how to spell foreign words in Korean *Hangeul* (한글).

Korea published 'Korean Orthography of Loanwords'[8] under the 'Framework Act on the National Language (국어기본법)'[9], and these guidelines govern all loanwords. The definition of a loanword, which is most widely accepted in Korea, is *a word that has assimilated into Korean and used as a Korean word*. As of October 31 in 2018, there were 59,210 words included in the Korean loanword orthography examples on the webpage of National Institute of Korean Language, and 47 words out of them were from Slovene. These loanword examples are mostly proper nouns such as personal and place names, and this is the same for Slovene words. Although the guidelines are called Korean orthography of loanwords, in reality, there are rather few real loanwords, which were assimilated into Korean and used as Korean words, such as 셔츠 'shirts', 아파트 'apartments', 라디오 'radio', 미팅 'meeting' etc. Lee Sang-gyu (2011, pp. 140-141) says that it is appropriate to call the current Korean orthography of loanwords as Korean orthography of foreign words in a strict sense as these guidelines are used to spell not only loanwords but also foreign words. With the development of technology and internet, the whole world became a global village, and proper nouns such as place names and personal names are nowadays exposed in real time and sometimes unexpectedly. Such words often enter written or spoken language for a single use, however, after several times they may become assimilated into Korean. It is not possible to determine whether foreign words with Korean spelling are loanwords or not as it takes time for a foreign word to get assimilated into Korean and become a loanword. This paper addresses Slovene words rather than Slovene loanwords, but the expressions Korean Orthography of Loanwords (KOL hereafter) and the Publication of

---

[8] This is our English translation of 외래어 표기법 because there is no official English name. These guidelines will be mentioned in this paper many time with the abbreviation KOL.

[9] The purpose of 'Framework Act on the National Language (국어기본법)' is to improve the quality of cultural life of the people and to contribute to the development of national culture by promoting creative thinking power of the people by promoting the use of the Korean language and setting the basis for development and conservation of Korean language, and this act covers how to promote the usage of Korean language, how to improve the Korean language ability of the people and how to popularize. This act is the foundation of 'Korean Orthography (한글 맞춤법)', 'Korean Standard Language Regulation (표준어 규정)', 'Korean Standard Pronunciation Rules (표준 발음법)', 'Korean Orthography of Loanwords (외래어 표기법)' and 'Revised Romanization of Korean (국어의 로마자 표기법)'.

Loanword Orthography Examples will be maintained because they are official names at present.

For now, there are no guidelines on how to spell Slovene words in Hangeul. There exist guidelines for Serbo-Croatian, the official language of former Yugoslavia, which included areas of the present day Serbia, Croatia, Bosnia and Herzegovina, Montenegro, North Macedonia, Kosovo, and Slovenia. Close examination of the Korean orthography and 47 Slovene words listed in the Loanword Orthography Examples showed that the words were spelled based on the rules for Serbo-Croatian. In principle, the orthography of personal names and place names from foreign origin should follow the provisions of Chapter 1, Chapter 2 and Chapter 3 of the KOL, however, personal names and place names from the languages which are not listed in KOL should be spelled based on their original pronunciation. In spite of these guidelines, in case of new foreign words from Slovene, people tend to follow the KOL for Serbo-Croatian probably because of the historical background and because Serbo-Croatian and Slovene belong to the group of South Slavic languages. However, Serbo-Croatian is not the official language of the Republic of Slovenia[10], and is different from Slovene, so KOL for Serbo-Croatian should not be accepted as appropriate for Slovene words. For example, Peter Prevc is spelled as 페테르 브레브츠 /peter prevʧ/ in the loanword orthography examples, but it is recommended to be spelled as 페테르 브레우쯔 /peter preuts/ because "v" in "Prevc" is pronounced as /u/ before consonants or in word final positions in Slovene. Sounds written as c /ts/ and č /tʃ/ in Slovene are different phonemic sounds. Another example is the case of Mitja Gasparini. His name in Korean is written as 미차 가스파리니 /miča gasparini/, and people call him 가스파 /gaspa/ although his name is 미탸 /mitja/. Besides, when we look into how Koreans use Slovene words for their communication, we can see that Koreans spell Bohinj in two ways: 보힌/bohin/ pronouncing nj as /n/ and 보히니/bohini/ pronouncing nj as [nɪ] on the internet and say the same in their conversation although 보힌/bohin/ is closer to the native pronunciation. All these examples lead to confusion. As there has been no discussion on this issue either in Korea or in Slovenia until now, this study would deal with the topic by providing information on Slovene alphabet and its pronunciation to avoid further confusion and to provide the foundation for detailed guidelines on Slovene loanwords in Korean.

In this study, we first introduce Korean Orthography of Loanwords (KOL) and its history in Section 2.1 and 2.2. We further describe KOL for Serbo-Croatian in Section 2.3, and current examples of Slovene words spelled in Korean in Section 2.4. Slovene alphabet and its pronunciation are described in Section 3.1 and 3.2. Our suggestion of detailed guidelines for KOL for Slovene is presented in Section 3.3, and propose amendments of some Slovene words spelled in Korean based on today's Slovene

---

[10] Lee Sang-gyu (이상규, 2011, p. 163) points out this matter. See, "Fundamental Law of languages" based on "foreign words notation" Analysis (국어 기본법에 근거한 외래어 표기법의 문제). The Korean Language and Literature (국어국문학회).C

alphabet and pronunciation in Section 3.4. In the Conclusion, we present a comparison table of the Slovene alphabet (abeceda) and Hangeul, and the Korean orthography examples of typical Slovene names, surnames and major city names as appendices.

2 Korean Orthography of Loanwords

2.1 Korean Orthography of Loanwords and its history

In South Korea, spellings of words from other languages in Hangeul are guided by the Korean Orthography of Loanwords, which was introduced on 7 January 1986 by the former Ministry of Culture and Education (문교부) of Republic of Korea. Now the National Institute of Korean Language (국립국어원) under the Ministry of Culture, Sports and Tourism (문화체육관광부) is responsible for these guidelines, modifications and supplementation.

Since September in 1991, The Joint Committee of the Government and Press on Loanwords Review (정부·언론 외래어 심의 공동위원회) [11] under the National Institute of Korean Language (국립국어원) has made several decisions and reviews on how to spell foreign words and loanwords in Korean, which is the base for the National Language Deliberation Council (국어심의회) to announce loanwords after deliberation. [12]

In the KOL, there are basic principles and detailed guidelines for spelling words from English, German, French, Spanish, Italian, Japanese, Chinese, Polish, Czech, Serbo-Croatian, Romanian, Hungarian, Swedish, Norwegian, Danish, Malay-Indonesian, Thai, Vietnamese, Portuguese, Dutch, and Russian. There is also a comparison table of alphabets of languages, and corresponding Hangeul. [13] Among 21 languages, four languages, i.e. Polish, Czech, Serbo-Croatian, and Russian, belong to Slavic languages. The KOL is trying to reflect the original local pronunciation (원지음) of each language as much as possible.

The orthography used in South and North Korea from around 1930s was based on the Draft for Standardized Korean Orthography (한글맞춤법 통일안, 1933) established by the Korean Language Society (조선어 학회, 현 한글 학회). A part of it is dedicated to the orthography of loanwords. In this part of the guidelines it is specified that the orthography of loanwords should be based on phonetic realization of the word origin, and that it is not allowed to use other alphabets or signs than Korean consonants

---

[11] As to the Joint Committee of the Government and Press on Loanwords Review, refer to the webpage of National Institute of Korean Language. https://www.korean.go.kr/popup/committee.do

[12] As to the deliberation on orthography of loanword, refer to article 4.4. Retrived from National Law Information center. http://www.law.go.kr/법령/국어심의회규정/(01977,19641110)

[13] But as for English, German and French, IPA and Hangeul comparison table is specified with detailed guidelines instead of individual language's alphabet and Hangeul.

and vowels. After several revisions, the current KOL was set up in 1986 in South Korea, and there have been some amendments until now. The brief history is in Table 1 below.[14]

Tabela 1: History of Korean orthography of loanwords

Institution	Year	Name of Transcription Guidelines
Korean Language Society (조선어 학회)	1933	Draft for Standardized Korean Orthography that covers Loanwords transcription (한글 맞춤법 통일안)
Korean Language Society (조선어 학회)	1940	Draft for Standardized Loanwords Transcription Guidelines (외래어 표기법 통일안)
Former Ministry of Culture and Education (문교부)	1948	Guidelines for transcribing adopted words from abroad[15] (들온말 적는 법)
Former Ministry of Culture and Education (문교부)	1958	Guidelines for transcribing the Roman alphabet in Korean (로마자의 한글화 표기법)
Former Ministry of Culture and Education (문교부)	1986	KOL, Notification No. 85-11 of the Ministry of Culture and Education (외래어 표기법, 문교부 고시 제 85-11 호)[16]
The National Institute of the Korean Language (국립국어연구원)	1992	Supplements for languages in the East European bloc, Notification No. 1992-31 of the Ministry of Culture (동구권 언어의 자모와 한글 대조표 제시, 문화부 고시 제 1992-31 호)[17]

[14] History of Orthography Guidelines of Loanwords, refer to Lee, Sang-gyu (2011, p. 149); Lee, Gyeong-suk (2016, pp. 111-136); Lim, Hong-bin (1996, pp. 28-29) and sources retrived from webpage of National Institute of Korean Language

https://www.korean.go.kr/niklintro2/20years05_01_02.jsp.

[15] Among the guidelines mentioned in Table 1, the 'Guidelines for transcribing adopted words from abroad (들온말 적는 법)' announced in 1948 were different from others in the sense that they allowed to spell foreign words based on their phonetic values. In order to represent foreign sounds which do not exist in Korean phonetic system, ancient Hangeul letters, i.e. ㅿ, ㅸ, ㆄ, and ㅹ, were reintroduced for specific foreign words. However, these guidelines were not maintained for a long time and their use was omitted in 1958, based on the opinion that the complexity impeded wide dissemination of these guidelines. See National Library of Korea

http://www.nl.go.kr/nl/search/search.jsp?all=on&topF1=title_author&kwd=들온말%20 적는%20 법.

[16] These guidelines included detailed rules for English, German, French, Spanish, Italian, Japanese and Chinese.

[17] The comparison tables of corresponding alphabet and Hangeul as well as detailed guidelines for Polish, Czech, Hungarian, Serbo-Croatian, and Romanian supplemented and announced.

Institution	Year	Name of Transcription Guidelines
The National Institute of the Korean Language (국립국어연구원)	1995	Supplements for languages in the North European block, Notification No. 1995-8 of the Ministry of Culture (북구권 언어의 자모와 한글 대조표 제시, 문화체육부 고시 제 1995-8 호)[18]
The National Institute of the Korean Language (국립국어원)[19]	2004	Supplements for languages in Southeast Asia, Notification No. 2004-11 of the Ministry of Culture and Tourism (동남아시아 3 개 언어 자모와 한글 대조표 제시, 문화관광부 고시 제 2004-11 호)[20]
The National Institute of the Korean Language (국립국어원)	2005	Supplements for Portuguese, Dutch and Russian, Notification No. 2005-32 of the Ministry of Culture and Tourism (포르투갈어, 네덜란드어, 러시아어 자모와 한글 대조표 제시, 문화관광부 고시 제 2005-32 호)
The National Institute of the Korean Language (국립국어원)	2014	Partial amendments of Korean orthography of loanwords 1986, Notification No. 2014-43 of the Ministry of Culture, Sports and Tourism (외래어 표기법 일부 개정안, 문화체육관광부 고시 제 2014-43 호)
National Institute of Korean Language (국립국어원)[21]	2017	Partial amendments of Korean orthography of loanwords 1986, Notification No. 2017-14 of the Ministry of Culture, Sports and Tourism (외래어 표기법 일부 개정안 문화체육관광부 고시 제 2017-14 호)

As can be seen in Table 1, the guidelines currently used are the KOL announced with Notification No. 85-11 of the Ministry of Culture and Education in 1986. After the announcement, the guidelines have been updated with supplements in additional languages. Besides, in case of mistakes, missing information, or inconsistency among language rules, amendments have been made to avoid confusion and to present correct language guidelines. We expect that such supplements for additional languages and partial amendments will be issued in the future as well.

Before the enactment of these rules, the Textbook Compilation Bureau under the Former Ministry of Culture and Education decided how to spell loanword and published

[18] A comparison table of corresponding alphabet and Hangeul as well as detailed guidelines for Swedish, Norwegian and Danish supplemented and announced.

[19] Only Korean name of the institute changed from 국립국어연구원 (The National Institute of the Korean Language) to 국립국어원 (The National Institute of the Korean Language) in 2004.

[20] A comparison table of corresponding alphabet and Hangeul as well as detailed guidelines for Malay-Indonesian, Thai and Vietnamese supplemented and announced.

[21] The English name of the institute changed from The National Institute of the Korean Language to National Institute of Korean Language in 2015.

Textbook Editing Materials (편수자료), proposing the related examples for each subject and those textbook editing materials were published by 1987.[22]

With the enactment of KOL, National Institute of Korean Language published *Loanwords Orthography Examples* with explanations in order to disseminate the guidelines consistently:[23]

- 1988 Publication of Loanwords Orthography Examples (general term, place names and personal names[24]).[25] 외래어 표기 용례집 (일반 용어, 지명, 인명)
- 1993 Publication of Loanwords Orthography Examples: place names and personal names in the East European bloc 외래어 표기 용례집 (동구권 인명 · 지명)
- 1995 Publication of Loanwords Orthography Examples: place names and personal names in the North European bloc 외래어 표기 용례집 (북구권 인명 · 지명)
- 2002 Publication of Loanwords Orthography Examples: place names 외래어 표기 용례집 (지명) Publication of Loanwords Orthography Examples: personal names 외래어 표기 용례집 (인명)
- 2004 Publication of Loanwords Orthography Examples for 3 languages in Southeast Asia: Malay-Indonesian, Thai and Vietnamese 동남아 3 개 언어 외래어 표기 용례집 (말레이인도네시아어, 타이어, 베트남어)
- 2005 Publication of Loanwords Orthography Examples: Portuguese, Dutch and Russian 외래어 표기 용례집 (포르투갈어, 네덜란드어, 러시아어)

2.2   The composition and basic principles of KOL

KOL consists of four chapters: Chapter 1 "The basic principles of orthography", Chapter 2 "A comparison table of orthography", Chapter 3 "The detailed guidelines", and Chapter 4 "The principles of spelling personal names and place names". In Chapter 1, there are five clauses as shown below[26]:

> **Clause 1.** Loanwords should be written with the 24 Korean letters currently in use.
> **Clause 2.** One phoneme of a loanword should principally be written with one Korean letter (grapheme).

---

[22] Lee, Gyeong-suk (2016, p. 109).

[23] Cf. https://www.korean.go.kr/niklintro2/20years05_01_02.jsp

[24] Authors use personal names and place names in this paper instead of proper noun as these are stipulated in the guidelines.

[25] It was published by the Research Institute of the Korean Language 국어연구소 which is former National Institute of Korean language.

[26] The guidelines of KOL was translated by authors.

Clause 3. A syllable coda, called batchim 받침[27], should be written written using ' ㄱ, ㄴ, ㅁ, ㄹ, ㅂ, ㅅ, ㅇ'.

Clause 4. No tense consonants should be used for spelling plosives.[28]

Clause 5. For loanwords in use, which need corrections, the existing common usage should be respected. The scope and examples should be decided separately.

In Chapter 2 of the KOL, comparison tables of alphabet in 18 languages including IPA transcription of Hangeul are presented. In Chapter 3, detailed guidelines are given for 21 separate languages and they deal with differences between Korean and target languages, specifying that we should write in accordance with the comparison table, following distinct features of each language. In Chapter 4, the spellings of personal names and place names are proposed in Clause 1, the spellings of words that originate in Asian languages are proposed in Clause 2, and the spellings of names of seas, islands, rivers, mountains, etc. are proposed in detail in Clause 3.

These principles and clauses show that KOL is under the guidelines of Korean Orthography and requires the spellings of loanwords according to Korean phonological system. There is an opinion that the comparison table of each language is of no use at all because the IPA and Hangeul comparison table could be used for all languages like English, German, and French work (Lee, 2016). But it is important to make the comparison table and detailed guidelines for each foreign language, so that all users of KOL conveniently check and use them because not all users have knowledge about the IPA symbols and not all languages indicate their pronunciation with IPA symbols in their dictionary, e.g. Slovene, Italian, and Spanish, etc. In addition, as KOL proposed to spell proper nouns according to the original local pronunciation of each language, it is more appropriate to make the comparison table for each language not to make it a useless document. Other studies (Cho & Park, 2015; Choi, 2014; Jang, 2002; Park, 1997) have shown that in case guidelines do not reflect the original local pronunciation fully, this causes confusion and eventually stops experts of each language as well as general public to use them adequately.

---

[27] Hanguel is a syllabic alphabet and has 5 different types of syllable structures, i.e. V (e.g. 오 'five'), CV (e.g. 소 'cow'), VC (e.g. 알 'egg'), CVC (e.g. 강 'river'), and CVCC (e.g. 닭 'chicken'). In the syllable structures of VC, CVC, and CVCC, the last consonant or the cluster of two consonants are called batchim, and are written at the bottom of the syllabic structure. According to the Korean Standard Pronunciation Rules, only 7 consont sound values are possible in the position of batchim. Those are ㄱ, ㄴ, ㄹ, ㅁ, ㅂ, ㅅ, and ㅇ.

[28] Korean has plain, aspirated and tense sounds for both plosives and affricates, and plain and tense sounds for fricatives. According to this clause, tense sounds, i.e. ㄲ, ㄸ and ㅃ are not allowed to use to transcribe/spell loanwords. Additionally, as to loanwards derived from English, German, French and Italian, it is not allowed to use tense sounds for affricates and fricatives, ⅰ.e. ㅉ and ㅆ are not allowed to use. For example, Paris 파리 not ＊빠리, service 서비스 not ＊써비스, Mozart 모차르트 not ＊모짜르트.

There are several controversial issues in KOL. To begin with, there is no clear distinction between foreign words and loanwords. Secondly, it is not clear whether proper nouns are the objects of KOL or not. Thirdly, there are conflicting opinions on whether transcription should be based on the pronunciation of a source language or Korean phonological system. Fourthly, there are many mistakes in loanwords' orthography examples, or else KOL is not followed well. Finally, different principles have been applied for loanwords of different origin. Words from English, German and French, for example, are spelled based on the IPA and Hangeul comparison table while words from the other languages are spelled based on their own alphabet and Hangeul comparison tables.

Therefore, we first propose to change the name of guidelines from 'Korean Orthography of Loanwords' to 'Korean Orthography of Foreign Words'. As mentioned above, it is almost impossible to know whether words from foreign languages could be assimilated into Korean or not in this early stage of spelling foreign words in Korean. If the name of guidelines is changed, the scope of KOl would be expanded and this could be a solution for arguments on whether proper nouns are scope of KOL or not. As for the principles of spelling loanwords, it should be acknowledged that a user of KOL is most likely Korean, and that a user can use it freely and conveniently only when they feel comfortable with the principles. Besides, the principles are ment only for spelling foreign words[29], and that they should be based on Korean phonological system while trying to reflect the pronunciation of a source language. Additionally, it seems that more government's efforts are needed to call attention from academic circles, publishing houses, mass media, and general public to obtain critical thinking and suggestions on the existing guidelines.[30] Lastly, we suggest that the guidelines including the IPA and Hangeul comparison tables for spellings of German and French words do not sufficiently represent pronunciations of words in a source language. Therefore, it would be appropriate to make their own alphabet and Hangeul comparison tables with detailed guidelines separately like it was done for other languages.

2.3   KOL for Serbo-Croatian

The guidelines with 72 examples for Serbo-Croatian were added to KOL in 1992, though in those times the Socialist Federal Republic of Yugoslavia has already collapsed. Both Slovene and Serbo-Croatian belong to the group of South Slavic languages. Officially, Slovene language has 25 letters (20 consonant and 5 vowel letters) and Serbo-Croatian 30 letters (25 consonant and 5 vowel letters), however, these two languages have some different pronunciation rules. As of January 2019, there were 269 loanword

---

[29] Yu Kyung-Sook. (2002, p. 17) says that these guidelines are not to spell the sounds of foreign languages to teach or learn the pronunciation of foreign languages but to present the orthography of foreign words within our phonological system for the Korean's consistency in use.

[30] Mistakes in Korean orthography examples of Slovene words are detected. See. footnote, 32-35.

orthography examples belonging to the section of Serbo-Croatian, with mainly proper nouns originating from languages in Slovenia, Croatia, Serbia, Montenegro, Kosovo, and even Bulgaria. Below are the detailed KOL guidelines for Serbo-Croatian. A comparision table of Serbo-Croatioan alphabet and Hangeul is in Table 2 below.

Table 2 : Comparison Table of Serbo-Croatian alphabet and Hangeul with examples[31]

	Hangeul		Orthography examples
	before V	before C, at the final	
Consonants			
b	ㅂ	브	bog 보그, drobnjak 드로브냐크, pogreb 포그레브
c	ㅊ	츠	cigara 치가라, novac 노바츠
č	ㅊ	치	čelik 첼리크, točka 토츠카, kolač 콜라치
ć, tj	ㅊ	치	nać 나치, sestrić 세스트리치
d	ㄷ	드	desno 데스노, drvo 드르보, medved 메드베드
dž	ㅈ	지	džep 제프, narudžba 나루지바
đ, dj	ㅈ	지	Đurađ 주라지
f	ㅍ	프	fasada 파사다, kifla 키플라, šaraf 샤라프
g	ㄱ	그	gost 고스트, dugme 두그메, krug 크루그
h	ㅎ	흐	hitan 히탄, šah 샤흐
k	ㅋ	ㄱ, 크	korist 코리스트, krug 크루그, jastuk 야스투크
l	ㄹ, ㄹㄹ	ㄹ	levo 레보, balkon 발콘, šal 샬
lj	ㄹ,리	ㄹ	ljeto 레토, pasulj 파술
m	ㅁ	ㅁ, 므	malo 말로, mnogo 므노고, osam 오삼
n	니+	ㄴ	nos 노스, banka 반카, loman 노만
nj	ㄴ	ㄴ	Njegoš 네고시, svibanj 스비반
p	ㅍ	ㅂ, 프	peta 페타, opština 옵슈티나, lep 레프
r	ㄹ	르	riba 리바, torba 토르바, mir 미르
s	ㅅ	스	sedam 세담, posle 포슬레, glas 글라스
š	시	슈, 시	šal 샬, vlasništvo 블라스니슈트보, broš 브로시
t	ㅌ	트	telo 텔로, ostrvo 오스트르보, put 푸트
v	ㅂ	브	vatra 바트라, olovka 올로브카, proliv 프롤리브
z	ㅈ	즈	zavoj 자보이, pozno 포즈노, obraz 오브라즈
ž	즈	주	žena 제나, izložba 이즐로주바, muž 무주

---

[31] National Institute of Korean Language. Retrived from
https://www.korean.go.kr/front/page/pageView.do?page_id=P000112&mn_id=97

Semivowel		
j	이※	pojas 포야스, zavoj 자보이, odjelo 오델로
Vowel		
a	아	bakar 바카르
e	에	cev 체브
i	이	dim 딤
o	오	molim 몰림
u	우	zubar 주바르

※ When a vowel follows after lj, nj, š, j, it is written together with the vowel as one syllable '리, 니, 시, 이'.

2.3.1 Detailed guidelines[32]

There are 5 clauses that payed our attention.

Clause 1. When *k, p* are a syllable coda position and appear before voiced consonants, they should get the epenthetic vowel 으 /ɯ/, but when these are placed before voiceless consonants, they should be transcribed as a syllable final consonant *받침*.

　　jastuk 야스투크, opština 옵슈티나

Clause 2. When *l* and *lj* in the middle of a word appear before vowels, they should be transcribed as *ㄹㄹ* /-ll-/, *ㄹ리* /-lli/.

　　kula 쿨라, Ljubljana 류블랴나

Clause 3. When initial consonant letter *m* is placed before *l, r, or n,* or else *m* appears in a middle position before *r*, it should be transcribed together with the following epenthetic vowel 으/ɯ/.

　　mlad 믈라드, mnogo 므노고, smrt 스므르트

Clause 4. When *š* stands before a consonant, it should be transcribed as 슈/ʃu/, while *š* in a final position becomes 시/ʃi/.

　　šljvovica 슐리보비차, Niš 니시

Clause 5. When consonants combine with *je*, they should get transcribed as 에/e/ and not 예/je/. Exception is a consonant letter *s*, which transcribes as 셰/ʃje/.

　　bjedro 베드로, sjedlo 셰들로

---

[32] Detailed guidelines have been translated by authors.

With the Comparison Table of alphabet transcription and the above guidelines, the original local pronunciation is reflected as close as possible with Hangeul and Korean orthography.

2.4   Current status of loanword orthography samples from Slovene

Anyone interested in how to spell loanwords in Hangeul can find loanword orthography examples on the website of the National Institute of Korean Language by searching the following keywords: Hangeul orthography (한글표기), original language orthography (원어표기), name of country (국명), etc.

With the keyword 슬로베니아 (Slovenia) searched for on Oct 31, 2018, the list on the web showed 47 Slovene words in Korean, all of them being proper nouns.

Table 3: Loanword orthography examples of Slovene words
from the website of the National Institute of Korean Language[33]

No.	Slovene word	Loanword orthography in Korean
1	Ljubljana	류블랴나
2	Maribor	마리보르
3	Slovenia[34]	슬로베니아
4	Trieste[35]	트리에스테
5	Postojna	포스토이나
6	Drnovšek, Janez	드르노브셰크, 야네즈
7	Janez Premoze[36]	야네즈 프레모제
8	Kučan, Milan	쿠찬, 밀란
9	Janša, Janez	얀샤, 야네즈
10	Danilo Türk	튀르크, 다닐로
11	Čeplak, Jolanda	체플라크, 욜란다
12	Zolnir, Urska (Urška Žolnir)	졸니르, 우르슈카
13	Cop, Iztok (Iztok Čop)	초프, 이즈토크
14	Spik, Luka (Luka Špik)	슈피크, 루카
15	Zbogar, Vasilij (Vasilij Žbogar)	주보가르, 바실리
16	Tokic, Bojan (Bojan Tokić)	토키치, 보얀
17	Kozmus, Primož	코즈무스, 프리모즈
18	Polavder, Lucija	폴라브데르, 루치야

---

[33] The table is mentioned in this paper without any correction. Retrieved October 31, 2018 from https://www.korean.go.kr/front/foreignSpell/foreignSpellList.do;front=D60B9400A217F31411BB2782218AC7AF?mn_id=96&pageIndex=1 accessed Oct. 31, 2018.

[34] Slovenia is the English spelling of Slovenija. The official name is Republika Slovenija.

[35] Trieste is a city in a Italian-Slovene bilingual area of today's Italy, Slovene name for it is Trst.

[36] Original spelling of this surname Premoze is Premože.

No.	Slovene word	Loanword orthography in Korean
19	Debevec, Rajmond	데베베츠, 라이몬트
20	Isakovic, Sara (Sara Isakovič)	이사코비치, 사라
21	Brežice	브레지체
22	Celje	첼레
23	Cerknica	체르크니차
24	Črnomelj	츠르노멜
25	Drava 강	드라바강
26	Idrija	이드리야
27	Istria 반도	이스트리아반도
28	Julijske 산맥	율리이스카산맥
29	Kočevje	코체베
30	Kranj	크란
31	Krško	크르슈코
32	Kupa 강[37]	쿠파강
33	Mura 강	무라강
34	Murska Sobota	무르스카소보타
35	Novo Mesto	노보메스토
36	Piran	피란
37	Katarina Srebotnik	스레보트니크, 카타리나
38	Ptuj	프투이
39	Sava 강	사바강
40	Škofja Loka	슈코퍄로카
41	Tolmin	톨민
42	Žižek, Slavoj	지제크, 슬라보이
43	Potrč, Marjetica	포트르치, 마레티차
44	Janković, Zoran	얀코비치, 조란
45	Pahor, Borut	파호르, 보루트
46	Bratušek, Alenka	브라투셰크, 알렌카
47	Prevc, Peter	프레브츠, 페테르

When we exam loanwords orthography examples of Slovene words in Table 3, we find discrepancies between the original Slovene pronunciation and words written in Korean. Moreover, some of these words on the list are not Slovene. These shortcomings of the table come from the fact that the orthographies of these examples were made based on the guidelines for Serbo-Croatian. Therefore, it is necessary to review Slovene examples.

---

[37] Kupa river is the Croatian name. In Slovenia, this river is called Kolpa river. The river runs on the border of two nations.

3 Slovene alphabet and pronunciation[38]

3.1 Slovene alphabet

The Slovene alphabet has 25 letters: A a, B b, C c, Č č, D d, E e, F f, G g, H h, I i, J j, K k, L l, M m, N n, O o, P p, R r, S s, Š š, T t, R r, V v, Z z, Ž ž. The letters q, w, x, y are sometimes found in Slovene texts but only in foreign names or borrowed expressions, e.g. New York, x žarki 'x rays'.

3.2 Slovene vowel system

Standard Slovene has seven long (stressed) vowels and six short vowels (either stressed or unstressed) written with letters *a, e, i, o,* and *u.* To denote the quantity and quality of a stressed vowel, superscript diacritics: ´, ˆ, ` are used but only in dictionaries and grammars. The Slovene vowel system is illustrated in Table 4 below.

Table 4: Slovene vowel system[39]

	Stressed		Unstressed short
	Long	Short	
High	í [íː] ú [úː]	ì [ì] ù [ù]	i [i] u [u]
high mid	é [éː] ó [óː]		
Mid		ə́ [ə́]	e [ə]
low mid	ê [ɛ́ː] ô [ɔ́ː]	è [ɛ̀] ò [ɔ̀]	e [ɛ] o [ɔ]
Low	á [àː]	à [à]	a [a]

Different from Slovene vowels, Korean vowels are pronounced without stress accent on any particular vowel. The KOL guidelines do not deal with the distinction between long and short vowels.

3.3 Slovene consonant system

Concerning consonants, voiceless obstruents /p, t, f, k, s, ʃ/ have their voiced counterparts which occur as individual letters in the orthography, e.g. *p–b, t–d, k–g, f-v, s-z, š-ž.* Voiceless obstruents *h* and *c* have their voiced allophonic counterparts *ɣ* and

---

[38] This section of the paper is based on description of Slovene phonetics in Toporišič (2004), Herrity (2000), and Lečič (2012). The purpose of this article is to shed light on how to write Slovene words in Korean in Korea and the objects are mostly proper nouns. Therefore, we deal with the phonology at the word level and do not deal with suffixes, plurals, prepositions, and inflection.

[39] Herrity (2000, p. 7). Corresponding IPA symbols are in square brackets. ´ and ` represent stressed vowels. : represents long vowels.

dz. There are voiced sonorants /m, n, r, l, j/. Additionally, *d* and *ž* are used sequentially for the pronunciation of [dʒ] in a limited number of loanwords, e.g. džem 'jam'. In terms of place and manner of articulation, the Slovene consonant system is illustrated by the following table.

Table 5: Slovene consonant system[40]

	labial	labio-dental	dental	alveo-palatal	palatal	velar
voiceless stops	p		t			k
voiced stops	b		d			g
voiceless fricatives		f	s	š [ʃ]		h [x]
voiced fricatives		v	z	ž [ʒ]		[ɣ] ✶
voiceless affricates			c [ts]	č [tʃ]		
voiced affricates			[dz] ✶	dž [dʒ]		[ŋ] ✶
Nasals	M		n			
Lateral			l			
Roll			r			
Glides					J	
Approximants	[w] ✶ [ʍ] ✶ [u̯] ✶					

3.3.1   Obstruents and their allophones

Obstruents appear, following phonological processes of final devoicing and regressive assimilation at the word level. In the first phonological process voiced obstruents devoice at the end of a word, e.g. *grad* 'castle'>[grat]*, rob* 'edge'>[rop]*, laž* 'lie'>[laʃ]*, Janez* male personal name>[Janes]. The case of /v/ will be explained later. When two obstruents occur next to each other in a word, regressive assimilation takes place. A voiced obstruent followed by a voiceless obstruent becomes voiceless, e.g. *sladko* 'sweet'>[slatko], *têžko* 'heavy'>[tɛʃko], *Iztok* male personal name>[istok], and a voiceless obstruent followed by a voiced obstruent becomes voiced, e.g. *kdo* 'who'>[gdo]*, glasba* 'music'>[glazba], *odločba* 'rule, order'>[odloʤba].

The voiced labio-dental /v/ has allophones while it is pronounced as [v] only before vowels or before syllabic /r/, e.g. *vas* [vas] 'village', *vôda* [vɔda] 'water', *vrt* [vərt] 'garden'. In other positions,

---

[40] Peter Herrity (2000, pp. 15-16). Those appearing in square brackets with ✶ are positional variants of phonemes and in square brackets are corresponding IPA symbols.

(a) the phoneme /v/ is realized as [u] when it appears at a word final after a vowel or when it appears as a syllable coda before a consonant, e.g. *živ* 'alive'>[*iu*]*, nov* 'new'>[nou]*, pravzaprav* 'as a matter of fact'>[prauzaprau], *ôvca* 'sheep'>[ouca]. The ending –ev is pronounced as [–əu], e.g. *cerkev* 'church'>[cerkəu].

(b) /v/ may be realized as a voiced glide [w] at the beginning of a word, before a voiced consonant or a sonorant. For example, *vzeti* 'to take'>[wzeti]*, vlak* 'train' >[wlak] or when it occurs between an obstruent/sonorant and a sonorant/obstruent, e.g. *odvzeti 'to take away'*>[odwzeti] or at the end of a word following a sonorant, e.g. *barv 'color'*>[barw].

(c) /v/ is pronounced as a voiceless bilabial approximant [ʍ] at the beginning of a word, before voiceless obstruents, e.g *vse* 'all'>[ʍse]*, vsebina* 'content'>[ʍsebina] or when it occurs between a voiced obstruent and a voiceless obstruent, e.g. *predvčerajšnjim* 'the day before yesterday'>[prɛtʍʧerajʃɲim][41].

3.3.2   Sonorants and their allophones

Sonorants /m, n, l, r, j/ are all voiced. /n/ has a nasal velar counterpart *[ŋ]* which is realized only before phonemes /k, g, h/, as in *banka* 'bank'>[baŋka] for example.

/l/ is pronounced as [l] when it appears at the beginning of a word or before /j/, e.g. *les* 'wood'> [les], /polje/ 'field'> [polje], and in personal names, e.g. [*pavəl*]*,* [*vilko*]*, [karəl*]. When /l/ appears between two vowels or after a consonant, it is also pronounced as [l], e.g. *telefon* 'telephone'>[telefon]*, glava* 'head'>[glava]*, mlad* 'young'>[mlat]. However, when it is at the end of a word or at the end of a syllable before a consonant, /l/ is normally pronounced as [u], e.g. *bolha* 'flea'>/bouha/*, solza* 'tear'>/souza/*, pekel* 'hell' >/pəkəu/.

/r/ is a dental roll and it can be pronounced as a single consonant [r] or syllabically as [ər], i.e. schwa followed by [r]. The *r* is pronounced syllabically as [ər] when it precedes a consonant at the beginning of a word, or occurs between two consonants, e.g. *smrt* 'death'>[smərt]*, vrt* 'garden'>*[vərt], prt 'table cloth'*>[pərt].

As for the glide /j/[42], when it precedes a vowel, it forms [ja], [je], [ji], [jo], [ju]. In case of consonant clusters /lj/ and /nj/, consonant sequences are pronounced as two distinct sounds at the beginning of a word when followed by a vowel or when in intervocalic position, e.g. *ljubezen 'love'*>[ljubezen]*, volja 'will'*>[volja]*, sanje 'dream'*>[sanje]. However, when they precede a consonant or occur in word final position, they are either pronounced as a simple [l] or [n] in a dialect or as a single

---

[41] Herrity (2000, p. 18) says that the –d-devoices before ʍ.

[42] The j is considered a semivowel in Korea, and counted among vowels as there is no consonant that distinguishes a palatal approximant.

palatalized consonant [l] or [ɲ] like *Kranj* city name >[Kraɲ/, *konj 'horse'* >[koɲ], *knjiga 'book'*>/kɲiga/.

3.4 Suggestions for detailed KOL guidelines KOL of Slovene words

On the basis of the above information on KOL, the Slovene alphabet, and pronunciation, we would like to suggest detailed guidelines for Korean orthography of Slovene words.

Clause 1. Voiced obstruents

1. When *b, d, g, z, ž* appear at word final position or before a voiceless consonant, write as 프 [pɯ], 트 [tɯ], 크 [kɯ], 스 [sɯ], 시 [si].

> pogreb 포그레프 'funeral',
> grad 그라트 'castle',
> Janez 야네스 male name,
> Iztok 이스토크 male name,
> Matjaž 마탸시 male name,
> bog 보크 'god',
>
> odpad 오트파트 'dump'

2. When *b* appears before a voiceless consonant, write as the final coda of a syllable ㅂ [b][43].

> občina 옵치나 'municipality'

3. When *b, d, g, z, ž* appear before a voiced obstruent, write as 브 [bɯ], 드 [dɯ], 그 [gɯ], 즈 [zɯ], 주 [zu].[44]

> dobre 도브레 'good',
> zadnji 자드니 'last'
> razlika, 라즐리카 'difference',
> možgani 모주가니 'brain'

Clause 2. Voiceless consonants

1. When *p, t, s, k, f, c, h* appear at a word final position or before a voiceless consonant, write as 프 [pɯ], 트 [tɯ], 스 [sɯ], 크 [kɯ], 프 [pɯ][45] , 쯔 [cɯ], 흐 [hɯ].

---

[43] Cf. foot note no. 25.

[44] As for the pronunciation of syllables *(*ㅈ, ㅉ, ㅊ + palatalized vowels ya, yeo, ye, yo, yu), it is provided to pronounce as a monophthong, e.g. 자, 짜, 차, etc. by clause 5 of the rules for Korean standard pronunciation (표준 발음법 제 5 항).

[45] There are no labio-dental sounds in Korean, therefore we use the same p sound for both p and f, and use the same b sound for both b and v to spell foreign words in Korean.

Kamnik 캄니크 'city name',
šef 셰프 'boss',
Ptuj 프투이 'city name',
kruh 크루흐 'bread',
pust 푸스트 'karnival,'
kovanec 코바네쯔 'coin',
vrh 버르흐 'peak'

2. When *k, s* and *č* precede a voiced consonant, write as 그 [gɯ], 즈 [zɯ], 주 [zu].

kdo 그도 'who',
glasba 글라즈바 'music',
odločba 오들로주바 'rule, order'

3. When *č* appears at a word final position or before a voiceless consonant, make a syllable with 이 [i].

plačnik 플라치니크 'payer',
točka 토치카 'point',
moč 모치 'power'

4. When *k* and *p* appear before a voiceless consonant, write as the final coda of syllable ㄱ [g], ㅂ [b].[46]

lekcija 렉찌야 'lesson',
oktober 옥토베르 'october',
čipka 칩카 'lace'

5. When *š* appears before a consonant, write as 슈 [ʃju], when *š* is the final consonant, transcribe as 시 [ʃi]. When *š* appears before a vowel *a, e, i, o, u,* transcribe as 샤 [ʃa] , 셰 [ʃe], 시 [ʃi] , 쇼 [ʃo], [ʃu] 슈 respectively.

številka 슈테빌카 'number',
Bekeš 베케시 family name,
šef 셰프 'boss'

Clause 3. Sonorants

1. When a non-initial *l* appears before a vowel, write as ㄹㄹ [-ll]. Wwhen it appears at a word final position or at the end of a syllable and before a consonant, write as 우 [u]. In names, write as ㄹ [l] before a consonant.

Izola 이졸라 city name,
Ljubljana 류블랴나 city name,
Pekel 페케우 cave name 'hell'

---

[46] Cf. foot note no. 43.

bolha 보우하 'flea',
Vilko 빌코 male name

2. When an initial consonant *m* appears before *l, r, n,* write as 므 [mɯ]. When m appears at the end of a word or before a consonant, write as the final coda of a syllable.

mlad 믈라트 'young',
mnogo 므노고 'many',
sejem 세옘 'fair'

3. When *n* appear before *k, g, h*, write as a final consonant of syllable o [ŋ].

banka 방카 'bank',
Angela 앙겔라 female name

4. When *r*[47] appears before a voiced consonant or at the end of a word, write as 르[rɯ]. But when it appears initially before a voiceless consonant or between two consonants, write as 어르 [ər].

torba 토르바 'bag',
mir 미르 'peace,'
smrt 스머르트 'death,'
vrt 버르트 'garden'

5. When *j* appears before a consonant in the word middle position or after a vowel in syllable coda, write as 이 [i]. But when *j* is combined with vowels *a, e, i, o, u* transcribe as 야 [ja], 예 [je], 이 [i], 요 [jo] or 유 [ju]. When *j* appears in a sequence with another consonant, write as one syllable. When *nj* and *lj* appear at word final, the *j* is not expressed in the Korean transcription.[48]

jagoda 야고다 'strawberry',
Velenje 벨레녜 city name,
Postojna 포스토이나 city name
Ptuj 프투이 city name,
Kranj 크란 city name,
prijatelj 프리야텔 'friend'

---

[47] There are no dental-roll sounds in Korean, therefore we use the same l sound for both l and r to spell foreign words in Korean.

[48] Although they are a single palatalized consonant [ʎ] or [ɲ], there is no Hangeul to reflect relative sounds. Therefore, we should use a consonant according to the guidelines as it is required to use only Hangeul to write foreign words in Korean.

Clause 4. Consonant *v*

When *v* appears before a vowel, transcribe as ㅂ [b][49]. When it appears at a word final position after a vowel or when *v* appears at the end of a syllable before a consonant, transcribe as 우 [u].

> vino 비노 'wine',
> nov 노우 'new,'
> vse 우세 'all',
> avto 아우토 'car'

Clause 5. Consonant sequence dž

When *d* is followed by *ž*, write as ㅈ [dʒ].

> džamija 자미야 'mosque'

3.5   Proposals for amendments of Korean orthography examples in case of Slovene words

Based on the Korean orthography of loanwords and the suggested details for Slovene, we would like to propose amendments for some words that are alreadz included in the orthography examples from Slovene.

In accordance with suggested clauses above,

1. Clause 1.1: Janez should be written as 야네스 and not *야네즈. Iztok should be written as 이스토크 and not *이즈토크. Finally, Primož should be written as 프리모시 and not *프리모즈.

2. Clause 3.3: Alenka should be written as 알렝카 and not *알렌카.

3. Clause 3.5: Celje should be written as 쩰레 and not *첼레. Also, Marjetica should be written as 마레티짜 and not *마레티차.[50]

4. Clause 4: Drnovšek should be written as 드르노우셰크 and not *드르노브셰크. Kočevje should be written as 코체우예 and not *코체베, and Polavder should be written as 폴라우데르 and not *폴라브데르.

In addition, we would like to suggest the use of the tense sound ㅉ in order to spell the *c* sound[51]. Although it is not allowed to use tense consonants for loanword

---

[49] cf. footnote no. 43.

[50] As to Slovene c sound, we suggest to use the tense sound ㅉ and the reason is explained.

[51] In a comparison table of Serbo-Croatian and Hangeul, for both c and č sounds, a Korean sound ㅊ [tʃ] is used because of the clause 4 of chapter 1 which does not allow tense sounds.

orthography, there were some exceptions with Malay·Indonesian, Thai, Vietnamese, and the confirmed loanwords from Japan already. Besides, there was an argument that this orthography matter with tense sounds should be reviewed for both Eastern European languages and Russian (Lee, 2016). In Slovene, the letter *c*, which is pronunced as [ts], has the same quality as Hangeul ㅉ, a tense consonant. Slovene words in question do not derive from English, German, French or Italian. There is a typical and popular women's name Mojca [moitsa] (모이짜) and it sounds very strange if we pronounce it as *모이차 [moitʃa], which is based on KOL for Sebro-Croatian. Therefore, considering the existence of one to one matching consonant and the original sound, it is desirable to accept the use of the tense consonant ㅉ [ts] for Korean orthography of Slovene words.

From the sociocultural point of view, in the 1980s when KOL was prepared and announced, it is understandable that one Korean alphabet was used for two different foreign alphabets. The first reason is that it was simple and convenient, and secondly, it was not easy for Koreans to distinguish the difference between two similar sounds: c and č. However, with an increasing exposure to foreign languages there is no problem to detect the difference between the two sounds and we feel rather uncomfortable to use the same spelling for two different sounds. In addition, considering the principle of orthography, and phoneticism, it is more proper to use the tense sound ㅉ for c.

Then, Lucija is 루찌야 rather than *루치야, Brežice becomes 브레지쩨 and not *브레지체, Celje is written as 쩰레 and not *첼레, Cerknica is 쩨르크니짜 and not *체르크니차 and Marjetica is 마레티짜 and not *마레티차 in order to reach the more correct pronunciation of Slovene names Korean.

4   Conclusion and proposal

The purpose of this paper is to introduce KOL (Korean Orthography Guidelines of Loanwords, 외래어 표기법), to review the current Korean orthography examples of Slovene words, and to shed light on the necessity to add up-to-date guidelines for Slovene to the KOL by proposing detailed guidelines with a comparison table of Slovene alphabet and Hangeul. There are several reasons for this. First, Korean orthography for Slovene words is composed based on the guidelines for Serbo-Croatian although there are differences between the two languages. Secondly, there are some Slovene words written wrongly in Korean orthography examples as well as on the internet. Such mistakes certainly lead to confusion. Finally, there is an endeavor to spell Slovene words correctly. Therefore, after introducing Korean Orthography Guidelines for Loanwords and their history, we looked in detail into the guidelines for Serbo-Croatian and Slovene alphabets, and Slovene pronunciation at a word level in order to review Korean orthography examples of Slovene words and to propose the detailed guidelines for Slovene. In the end, we added a comparison table of Slovene alphabet and Hangeul.

In addition, we proposed to change the name of the guidelines into "Korean Orthography of Foreign Words" as a solution to the arguments on the distinction between loanwords and foreign words, and its scope. Since the users of KOL are Koreans, and the guidelines have to make Koreans spell foreign words consistently, we suggested that guidelines be based firstly on Korean phonological system and orthography. At the same time, the pronunciation in the source language should also be reflected at its maximum. The table below is a comparison table of the Slovene alphabet and Hangeul as a summary of detailed guidelines for Slovene.

Table 6: Comparison table of Abeceda and Hangeul for Korean orthography of Slovene words

| Abeceda | Hangeul | | Orthography examples |
	before a vowel	before a consonant, final position	
Consonants			
B	ㅂ	ㅂ, 브, 트	bled 블레트, dobro 도브로, občina 옵치나, pogreb 포그레프
C	ㅉ	쯔	cesta 쩨스타, Logatec 로가테쯔
Č	ㅊ	치	čokolada 초콜라다, točka 토치카, Zupančič 주판치치
D	ㄷ	드, 트	delo 델로, vrednost 브레드노스트, grad 그라트
F	ㅍ	프	fakulteta 파쿨테타, šef 셰프
G	ㄱ	그, 크	gibanica 기바니짜, igra 이그라, bog 보크
H	ㅎ	흐	hrana 흐라나, uspeh 우스페흐
J	이	※	jagoda 야고다, postojna 포스토이나, kranj 크란
K	ㅋ	ㄱ, 그, 크	kava 카바, akcija 악찌야, kdo 그도, Kamnik 캄니크
L	ㄹ, ㄹㄹ	우, ㄹ	levo 레보, ljubljana 류블랴나, bolha 보우하, hotel 호텔, Pekel 페케우
M	ㅁ	므, ㅁ	moški 모슈키, mnog 므노그, sejem 세옘
N	ㄴ	ㅇ, ㄴ	narava 나라바, banka 방카, zakon 자콘
P	ㅍ	ㅂ, 프	pošta 포슈타, čipka 칩카, lep 레프
R	ㄹ	어르, 르	roza 로자, torba 토르바, smrt 스머르트, sir 시르
S	ㅅ	즈, 스	Slovenija 슬로베니야, desno 데스노, glasba 글라즈바, avtobus 아우토부스

Abeceda	Hangeul		Orthography examples
	before a vowel	before a consonant, final position	
Š	ㅅ	슈, 시	šef 셰프, številka 슈테빌카, Bekeš 베케시
T	ㅌ	트	Tolmin 톨민, avto 아우토, list 리스트
V	ㅂ	우	voda 보다, nov 노우
Z	ㅈ	스	zvezda 즈베즈다, Iztok 이스토크, Janez 야네스
Ž	ㅈ	주, 시	žena 제나, dežnik 데주니크, mož 모시
Vowels			
A	아		ananas 아나나스
E	에		evro 에우로
I	이		ime 이메
O	오		okno 오크노
U	우		ura 우라

※ When a vowel follows after lj, nj, š, j, it is written together with the vowel as one syllable '리, 니, 시, 이', and refer to the detailed guidelines.

We followed the Korean phonological and orthography system except for the sound of Slovene letter *c*, and propose to use a tense sound ㅉ for *c* instead of ㅊ although it is not allowed to use a tense sound in present KOL. The reasons are, first, that both *c* and ㅉ have same IPA sound [ts], secondly, that we can reflect the original pronunciation at its maximum, thirdly, that there were already exceptions with loanwords from some other languages, and lastly, that there is a movement that the guidelines should be reviewed in case of tense sounds for languages from the Eastern European countries.

There are more and more exchanges between Korea and Slovenia nowadays, and the popularity of Korean Studies at the Department of Asian studies at the Faculty of Arts, University of Ljubljana is increasing. This suggests that Korean orthography of Slovene is indispensable not only to Koreans who want to share their work and experience in Slovenia, but also to Slovenes who want to know how to write their names in Korean when they start studying Korean and other proper nouns later on to introduce their culture and society to Korean people in Korea. This study is the first step to discuss how to spell Slovene words in Korean, and to begin an open discussion on this issue. We are aware that further studies to support our proposal are needed. It would be of our best satisfaction if this research could provide those interested in KOL with the necessary information and call attention for the need of Korean Orthography for Slovene words.

In appendices, please find Korean spellings of the most frequent names, surnames, and largest cities and towns in Slovenia.

References

Cho, N.-S., & Park, S.-B. (2015). A study on the orthographic representation of Russian loanwords in Korean (러시아어 한글 표기법에 대한 제안). *Issledovanija po slavjanskim jazykam, 20*(2)*,* 2015.10, 197-213.

Choi, J. Y. (2014). Study on Korean Rules of Spelling of Russian Loan Words (러시아어의 한글표기실태 - 학술 논문집을 중심으로). *Korean Journal of Russian Language and Literature, 26*(4), 2014.12, 139-164.

Greenberg, M. L. (2006). *A Short Reference Grammar of Standard Slovene.* Durham, North Carolina: SEELRC.

Herrity, P. (2000). *Slovene: A Comprehensive Grammar.* London, New York: Routledge

I, G. (1935). Oeraeeo Pyogibeop-e daehaya (외래어 표기법에 대하야). *HAN-GEUL 3*(6)*,* 1935.8, 357-357.

Jang, H.-U. (2002). Problèmes de la notation actuelle du français en coréen et quelques proposition (불어의 한글 표기상 문제점과 개선안). *Societe Coreenne d'Enseignement de Langue et Litterature Francaises, 13*, 2002.5, 1-17.

Jung, J.-Y. (2017). Hangeul Notation on Dutch Words in Korean (외래어 표기법 속네덜란드어의 한글 표기법). *HAN-GEUL, (317)*, 2017.9, 211-238.

Kim, J.-W. (2008). Korean Orthography of Loanwords and Romanization of Korean From the Viewpoint of Translation (번역의 관점에서 본 외래어 표기법과 로마자 표기법). *The Journal of Translation Studies, 9*(2), 2008.06, 67-93.

Lečič, R. (2012). *Basic grammar of the slovene language, language manual*. Cerkno: Gaya d.o.o.

Lee, G.-S. (2016). The Transition Process and Acceptance of 'Korean Orthography of Loanwords' by Ministry of Education: With Emphasis on The Earlier Draft Plans and Editing. Materials of Ministry of Education before Enactment of Korean Orthography of Loanwords 문교부의 '외래어 표기법' 변천 과정과 수용 양상 – 외래어 표기법 (1986) 제정이전의 문교부안과 편수자료를 중심으로). *Korean Studies Quarterly, 39(1)*, 2016.3, 107-140.

Lee, H. H. B. (1989). *Korean Grammar*. Oxford: Oxford university press.

Lee, S.-G. (2011). Fundamental Law of languages based on "foreign words notation" Analysis. (국어기본법 근거한 외래어 표기법의 문제). *The Korean Language and Literature, (158)*, 2011.8, 135-181.

Park, K. Ja. (1997). Zur orthographischen Umsetzung des Deutschen ins Koreanische (독일어의한글표기에 관하여), *Deutsch als Fremdsprache in Korea, 2*, 1997.6, 294-316.

Toporišič, J. (2004). *Slovenska slovenica.* Maribor: Založba Obzorja d.d.

Yu, K.-S. (2002). Standardizing transliteration of Italian proper names for Korean library catalogs and bibliographic lists 이탈리아인명, 지명의 한글표기. *The Journal of Humanities, 5*, 2002.2. 15-43.

Ziraldo, I., & Krofič, J. (2014). Problems of Hangeul Romanization into Slovene. *Acta Linguistica Asiatica, 3*(3), 71-84. https://doi.org/10.4312/ala.3.3.71-84

National Institute of Korean Language 국립국어원. (n.d.) Retrieved from https://www.korean.go.kr/

National Library of Korea 국립중앙도서관. (n.d.) Retrieved from
http://www.nl.go.kr/nl/search/search.jsp?all=on&topF1=title_author&kwd=들온말 적는 법

National Law Information center 국가법령정보센터. (n.d.). Retrieved from
http://www.law.go.kr/법령/국어심의회규정/(01977,19641110)

Slovene alphabet and pronunciation (n.d.). Retrieved from http://slovlit.ff.uni-lj.si/sft/alphabet.htm

The most frequent Slovene first names (2018). Retrieved from
https://www.stat.si/ImenaRojstva/sl/FirstNames

The most frequent Slovene family names (2018). Retrieved from
https://www.stat.si/ImenaRojstva/sl/FamiliyNames

Največja slovenska mesta po velikosti [The biggest cities in Slovenia] (2007-2019). Retrieved
from http://kraji.eu/thematic_page/najvecja_slovenska_mesta_po_velikosti/slovenija/slo

Abbreviation and Symbols

C   consonant
IPA   International Phonetic Alphabet
KOL   Korean Orthography of Loanword
V   vowel
>/ /   read as
∗   incorrect form

Appendix 1

The most frequent Slovene personal names spelled in Korean based on our proposal to spell in Hangeul[52]

Male				Female			
No.	Names	Number of names	Names spelled in Korean	No.	Names	Number of names	Names spelled in Koran
1	Franc	23.554	프란쯔	1	Marija	55.190	마리야
2	Janez	20.955	야네스	2	Ana	25.463	아나
3	Ivan	17.406	이반	3	Maja	13.445	마야
4	Anton	17.228	안톤	4	Irena	12.099	이레나
5	Marko	17.149	마르코	5	Mojca	11.302	모이짜
6	Andrej	16.745	안드레이	6	Mateja	10.370	마테야
7	Jožef	15.198	요제프	7	Nina	10.257	니나
8	Jože	14.063	요제	8	Nataša	10.035	나타샤
9	Luka	12.508	루카	9	Andreja	9.350	안드레야
10	Peter	12.220	페테르	10	Barbara	9.325	바르바라

Appendix 2

The most frequent Slovene family names spelled in Korean based on our proposal to spell in Hangeul[53]

No.	Family names	Number of family names	Family names spelled in Korean
1	Novak	10.999	노바크
2	Horvat	9.496	호르바트
3	Kovačič	5.544	코바치치
4	Krajnc	5.476	크라인쯔
5	Zupančič	4.999	주판치치
6	Potočnik	4.689	포토치니크
7	Kovač	4.623	코바치
8	Mlakar	3.899	믈라카르
9	Kos	3.838	코스
10	Vidmar	3.809	비드마르

---

[52] Source retrieved January 1, 2018 from SURS https://www.stat.si/ImenaRojstva/sl/FirstNames
[53] Source retrieved January 1, 2018 from SURS https://www.stat.si/ImenaRojstva/sl/FamiliyNames

Appendix 3

Slovene city names spelled in Korean based on our proposal to spell in Hangeul [54]

No.	City names	Population	City names already in orthography examples	City names which are not in orthography examples and spelled in Korean
1	Ljubljana	280.940	류블랴나	
2	Maribor	95.171	마리보르	
3	Celje	37.520	첼레	쩰레[55]
4	Kranj	36.874	크란	
5	Velenje	25.456		벨레네
6	Koper	24.996		코페르
7	Novo mesto	23.341	노보메스토	
8	Ptuj	18.164	프투이	
9	Trbovlje	15.163		터르보울레
10	Kamnik	13.644		캄니크
11	Jesenice	13.440		예세니쩨
12	Nova Gorica	13.178		노바고리짜
13	Domžale	12.406		돔잘레
14	Škofja Loka	11.969	슈코퍄로카	
15	Murska Sobota	11.614	무르스카소보타	
16	Izola	11.223		이졸라
17	Postojna	9.183	포스토이나	
18	Logatec	8.942		로가테쯔
19	Kočevje	8.672	코체볘	코체우예[56]
20	Vrhnika	8.413		버르흐니카

[54] Source retrieved January 1, 2018 from
http://kraji.eu/thematic_page/najvecja_slovenska_mesta_po_velikosti/slovenija/slo

[55] We propose to amend the Korean orthography of this word from 첼레 into 쩰레.

[56] We propose to amend the Korean orthography of this word from 코체볘 into 코체우예.

SOGLASNIKI STANDARDNE KITAJŠČINE V OKVIRU SLOVENSKEGA GLASOVNEGA SISTEMA[*]

Mateja PETROVČIČ
Univerza v Ljubljani, Slovenija
mateja.petrovcic@ff.uni-lj.si

Abstract

This paper discusses the pronunciation of consonants of Standard Chinese from the perspective of native speakers of Slovenian. Slovenian distinguishes 21 consonants, whereas Chinese distinguishes 22. Since there are differences in the phonetic features of the two languages in question, this paper examines the strategies of how to map Chinese phonemes to Slovenian sounds. In this process, the principle of rationality plays an important role to avoid unnecessary complications. To achieve this goal, we first outline the categories of Slovenian and Chinese consonants, compare them using the International phonetic alphabet, and propose equivalents for the Slovenian reader. Guidelines for the suggested pronunciation are finally summarized in a short passage.

Keywords: Chinese; Hanyu pinyin; consonants; Slovenian; sound system

Povzetek

Članek obravnava soglasnike standardne kitajščine z vidika govorca slovenskega jezika. Slovenščina ima 21 pomensko razločevalnih soglasnikov, kitajščina enega več. Ker se fonetične značilnosti soglasnikov med jezikoma razlikujejo, prispevek prouči možnosti, kako s slovenskimi glasovi predstaviti kitajske foneme, pri čemer upoštevamo tudi načelo gospodarnosti. V ta namen najprej podamo oris slovenskih in kitajskih soglasnikov, jih primerjamo s pomočjo mednarodne fonetične abecede, predlagamo ustreznice in predlagane rešitve strnemo v krajše navodilo, kako naj slovenski bralec izgovarja kitajske soglasnike.

Ključne besede: kitajščina; pinjin; soglasniki; slovenščina; glasovni sistem

---

1   Uvod

Glasovni sistemi različnih jezikov se med seboj razlikujejo, pri čemer prihaja do razlik med razvrstitvijo in vlogo posameznih prvin glasovnega sistema. S fonološkega vidika je pomembno, da razumemo, katere foneme imamo na razpolago v določenem jeziku, saj so to najmanjše pomensko razločevalne enote. Zavedati se moramo, da že znotraj istega jezika prihaja do odstopanj med izgovarjavo in zapisom besed. Množice fonemov ne moremo enostavno preslikati v množico grafemov. Pri nekaterih jezikih je odstopanj med izgovarjavo in zapisom besed več, pri drugih manj. Zanima nas, kaj se dogaja pri vstopanju kitajskih besed v slovenski jezik.

Ko besede iz tujega jezika vstopajo v slovenski jezik, jih pogosto najprej zasledimo v pisni obliki. Grafeme interpretiramo v okviru slovenskih glasov, pri čemer upoštevamo nezavedne glasovne procese svojega jezika. Za lažje razumevanje navedimo primer iz angleščine. Če nismo posebej šolani v izgovarjavi angleškega jezika, bomo v angleško-slovenskem slovarju sicer izvedeli, da besedo 'pick' preberemo kot [pɪk] in 'pig' kot [pɪg], vendar bomo obe besedi *samodejno* izgovorili kot [pik], saj v slovenščini velja, da črko 'g' v izglasju izgovarjamo kot glas [k] (Slovenski pravopis, 2001, § 1069). Enako velja za vse črke za zveneče nezvočnike – izgovarjamo jih nezveneče.

To ne pomeni, da na koncu besede nismo zmožni izgovoriti zvenečega nezvočnika in ustrezno ločiti [pɪg] od [pɪk], temveč da moramo biti pri učenju angleščine pozorni na to razliko. Gotovo bi tudi znali ustrezno ponoviti izvorno izgovarjavo avtomobilskih znamk 'Peugeot' [pø:ˈʒo] ali 'Renault' [ʁəno], vendar dobita ti besedi v okviru slovenskih glasov podobo [pežó] in [renó] (gl. FRAN). Na primeru teh dveh besed vidimo, da zaokroženih glasov ne izgovarjamo zaokroženo, francoski grlni /r/ pa zamenjamo s slovenskim jezičnikom /r/. Podobno mehkega hrvaškega ali srbskega fonema /ć/ ne izgovarjamo mehko, temveč enako kot fonem /č/ (npr. Andrić [andrič] in Poreč [poreč]). Med značilnosti slovenščine sodi tudi zakonitost, da pred samoglasniki na začetku besede ne izgovarjamo dvoustničnega u̯, temveč zobноustnični [v]. Tako angleškega Walden ne preberemo kot [ˈwɔ:ldən], temveč kot [vóldən].[1]

 Razumljivo je, da v okviru slovenščine tuje glasove zamenjujemo z najbližjimi slovenskimi knjižnimi glasovi, pri čemer opuščamo vse vrste modifikacij, na primer pri mehčanih in mehkih soglasnikih, posebne vrste zapornikov in podobno (prim. Slovenski pravopis, Pravila, § 222).

Cilj tega prispevka je torej najti slovenske ustreznice za izgovarjavo grafemov za kitajske soglasnike, na katere naletimo v sistemu *Hanyu pinyin* ali poslovenjeno pinjin. Izraz 'pinjin' je občno ime za točno ta sistem, ki ga je Mednarodna organizacija za standardizacijo sprejela leta 1982, Združeni narodi pa leta 1986. Poleg sistema *Hanyu*

---

[1] Za ostale podrobnosti o fonemu /v/ gl. § 640–645.

pinyin namreč obstaja tudi *Tongyong pinyin*, ki se od prvega razlikuje v določenih podrobnostih.

Pri iskanju izgovornih rešitev za kitajske knjižne glasove se skušamo v slovenščini čim bolj približati izvirnemu izgovoru, pri čemer imamo pri presoji v mislih predvsem fonološki vidik. Najpomembneje je, da ohranimo pomensko razločevalne enote. Kjer bi stremenje k preveč dosledni glasovni podobi vodilo do komplikacij, se odločimo za enostavnejše rešitve, saj že v izhodišču vemo, da povsem identične glasovne podobe zaradi omejitev slovenskega glasovnega sistema ne bomo mogli doseči (gl. Slika 1). Slovenski govorec bo torej kitajske zloge v okviru slovenščine izgovarjal *po slovensko*, med komuniciranjem v kitajščini pa *po kitajsko*.
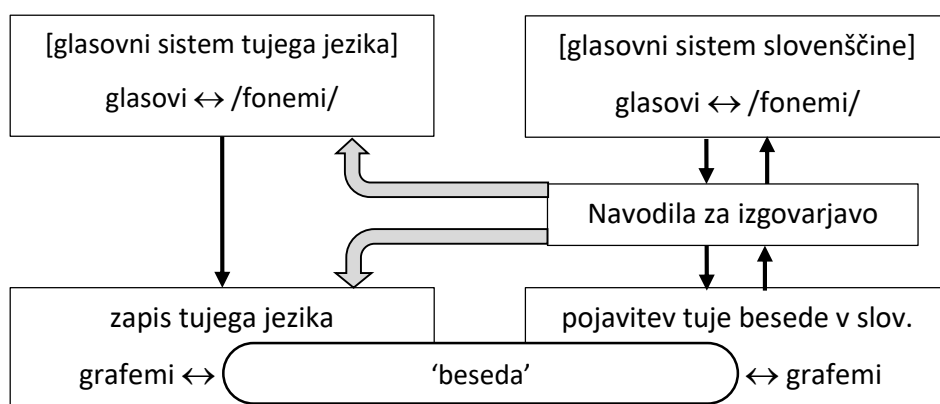
Slika 1: Konceptualni prikaz iskanja rešitev kitajsko-slovenskih ustreznic

Poznavanje glasovnega ustroja materinega jezika nam tudi pomaga razumeti in predvideti, kje bodo pri učenju tujega jezika nastopile težave, saj se moramo v tem procesu spoznati z glasovi in značilnostmi ciljnega jezika, ki se fonetično in fonološko razlikujejo od materinščine. S tega vidika imajo govorci različnih jezikov različne težave pri usvajanju glasovne podobe kitajščine. Poleg aspiracije (pridihnjenosti) imajo govorci angleščine na primer težave z zlitniki (Defense Language Institute, 1974; Comparison of English and Mandarin (Segmentals), 2019), nam povzročajo težave retrofleksni in nebni glasovi, govorcem češčine pa poleg tega še mehkonebni nesičnik pripornik (Švarný & Uher, 2001; Uher, Jin, & Slaměníková, 2016). Govorci jezikov, ki ne dopuščajo soglasnikov v izglasju, na primer Telugu (Sailaja, 1999), imajo težave z nosnikoma [n] in [ŋ]. S slednjim imamo težave tudi govorci slovenščine, saj mehkonebnojezični glas [ŋ] izgovarjamo le na stičišču glasov [n] in [k/g/h] (Anka, Angela, Anhovo), ne pa tudi v izglasju.

2   Primerjava soglasnikov v knjižni slovenščini in standardni kitajščini

2.1   Oris soglasnikov v slovenščini

Slovenski knjižni jezik ima 29 pomensko razločevalnih glasovnih enot (fonemov), od tega 8 samoglasnikov in 21 soglasnikov. Nam dobro poznana delitev soglasnikov je:

- zvočniki (/m/, /n/, /r/, /l/, /v/, /j/)
- nezvočniki
    - zveneči (/b/, /d/, /g/, /z/, /ž/, /dž/)
    - nezveneči (/p/, /t/, /k/, /s/, /š/, /č/, /c/, /f/, /h/)[2]

Razmerja med glasovi in črkami so natančneje prikazana v § 1070, razmerja med črkami in glasovi pa v § 1069. Zaradi razprave, ki sledi na naslednjih straneh, še enkrat poudarimo, da se *glas*, *fonem* in *grafem* ne prekrivajo v celoti. Izpostavimo nekaj zgledov iz § 1070:

Glas [g] zapisujemo s črko:

- 'g' (glagol, Bogdan, rog doni) in
- 'k' pred zvenečim nezvočnikom (kdaj, k domu, vsak dan).

Glas [k] zapisujemo s črko:

- 'k' (krik, kura, pakt) in
- 'g' pred nezvenečim nezvočnikom (bogca), v izglasju besede pred premorom (rog), pa tudi pred samoglasnikom ali zvočnikom naslednje besede (rog odmeva, rog joče)

Torej črka 'g' zapisuje:

- glas [g] (glagol, Bogdan, rog doni) in
- glas [k] pred nezvenečim nezvočnikom (bogca, rog poje), v izglasju besede pred premorom (rog), pa tudi pred samoglasnikom ali zvočnikom naslednje besede (rog odmeva, rog joče)

Sorodno temu črka 'k' zapisuje:

- glas [k] (krik, kura, pakt, k tebi, pek odhaja, prek jezera) in
- glas [g] pred zvenečim nezvočnikom (kdaj, k bratu, vsak dan).

Glas [dž], sicer vizualno dve enoti, zapisujemo:

- s črkama 'dž' (džungla),
- s črko 'č' pred zvenečim nezvočnikom (odločba, proč daj) in
- s črko 'c' pred zvenečima šumevcema ž in dž (stric Žane, stric Džon).

---

[2] Zvočnike si zapomnimo z izrazom *mlinarjev*, zveneče nezvočnike z izrazom *Božo gode džez*, nezveneče nezvočnike pa z izrazom *ta suhi škafec pušča*.

2.2 Oris soglasnikov v kitajščini

Standardna kitajščina ima 22 soglasnikov, ki so glede na mesto artikulacije razporejeni v naslednje skupine:

- dvoustnični /b/, /p/, /m/
- zobnoustnični /f/[3]
- jezičnokonični /d/, /t/, /n/, /l/, /z/, /c/, /s/
- retrofleksni /zh/, /ch/, /sh/, /r/
- jezično-ploskovni /j/, /q/, /x/
- mehkonebni /g/, /k/, /h/
- mehkonebni nosnik /ŋ/

Vsi soglasniki razen mehkonebnega nosnika /ŋ/ stojijo na začetku zlogov, nosnik /n/ stoji na začetku ali koncu zloga, mehkonebni nosnik /ŋ/, ki ga zapisujemo s črkama 'ng', pa stoji le na koncu zloga.

Med grafemi pinjina sta še črki w in y, ki služita kot vizualno sredstvo za ohranjanje hiata v zlogih tipa GV (drsnik+samoglasnik) (Golob & Petrovčič, 2018, str. 74). Laično gledano bi rekli, da sta to soglasnika, vendar je njuna raba precej bolj kompleksna.

3 Fonetični in fonemski zapisi soglasnikov obeh jezikov

Po številu soglasnikov sta slovenščina in kitajščina zelo primerljiva jezika. Oba imata 21 soglasnikov, pri čemer se kitajski dvaindvajseti soglasnik – mehkonebni nosnik /ŋ/ – razlikuje od ostalih v tem, da se pojavi le v izglasju.

Da bi lahko proučili oba glasovna sistema, smo ju morali najprej postaviti na skupni imenovalec, saj je nemogoče primerjati dve tako različni klasifikaciji. Kot orodje za ta korak nam je služila Mednarodna fonetična abeceda (IPA).

Namen **ozkega fonetičnega zapisa** je zapis glasov do take mere, da ohranimo natančno izgovarjavo. Za to uporabljamo oglate oklepaje. Ozki fonetični zapis je laiku težko berljiv. Za boljšo predstavo, kako so zapisane nianse med glasovi, priporočamo ogled naslednjih strani:

- J. Dowse (2018). IPA *Charts with Audio*.
  Dosegljivo na https://jbdowse.com/ipa/
- M. Deroń (2018). *IPA 2018 i-charts*. Dosegljivo na
  https://linguistics.ucla.edu/people/keating/IPA/inter_chart_2018/IPA_2018.html

---

[3] *Bopomofo*, tajvanski glasovni zapis kitajščine, je dobil svoje ime po prvih štirih elementih tega sistema.

Široki fonetični zapis zajame zgolj najbolj očitne elemente fonetičnega zapisa in prehaja v fonemski zapis. Za **fonemski zapis** običajno uporabljamo poševnice, pri tem pa sledimo načelu, kako so določeni glasovi v danem jeziku kategorizirani. Meja med fonetičnim in fonemskim zapisom ni vedno jasna in je stvar dogovora. Poglejmo si zapise angleške besede 'red' (rdeč) in kitajske besede 'gua' (buča):

1. podrobni fonetični zapis: [ɹʷɛd][4] … [kʷaa]
2. osnovni fonetični zapis: [ɹʷɛd] … [kʷa]
3. fonetično-fonemski zapis: [ɹɛd] … [kwa]
4. fonemski zapis: /red/[5] … /gua/[6]

Na tem mestu si oglejmo primerjavo slovenskih in kitajskih soglasnikov po tretji različici zapisa (Tabela 1 – *zaradi boljšega pregleda se celotna tabela nahaja na naslednji strani*).

Iz distribucije glasov opazimo, da fonetične razlike med jezikoma niso tako velike. Nosniki, nesični priporniki in stranski drsniki se popolnoma ujemajo. Pri slovenskih zapornikih in sičnikih zlitnikih je pomensko razločevalne narave zvenečnost, pri kitajskih pa aspiracija. Tudi distribucija sičnikov pripornikov je primerljiva s sičniki zlitniki, čeprav obstajajo navidezno večja odstopanja.

Do največjih razhajanj prihaja pri opisu glasu, ki ga v pinjinu zapisujemo s črko 'r'. Li in Thompson (1981, str. 5) navajata, da je to "zveneči retrofleks [ɹ]", pri čemer je najbrž mišljen zapis [ɻ] in ne [ɹ], ki se uporablja za zveneči zobnodlesnični drsnik. Duanmu (2007, str. 24) meni, da gre za retrofleksni sičnik pripornik [ʐ], obenem pa navaja, da starejše študije ta glas zapisujejo kot [r]. Tudi Huang (2018) uporablja za zapis [ʐ] ali [ɻ]. Po pregledu strokovne literature smo se odločili za zapis [ʐ], saj se največ študij nagiba k tej rešitvi.

---

[4] [ɹ] označuje alveolarni aproksimant, podčrtaj prikazuje, da je glas izgovorjen bolj zadaj v ustih kot običajno, [ʷ] pa signalizira labializacijo oz. zaokroženost ustnic.

[5] Vir: lingvo.info

[6] Vir: Duanmu (2007, str. xi)

Tabela 1: Slovenski in kitajski soglasniki ter njihove vrednosti v Mednarodni fonetični abecedi

mesto artik.	izdišni soglasniki													
	ustnični				koronalni				nebni				grlni	
	dvo-ustnični		zobno-ustnični		dlesnični		zadlesnični		retrofleksni		nebni		mehko-nebni	
način artik.	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN
nosnik		m				n								ŋ
		m				n								ŋ
zapornik	p	b			t	d							k	g
	p				t								k	
	pʰ				tʰ								kʰ	
sičnik zlitnik					ts	dz	t͡ʃ	d͡ʒ						
					ts				ʈʂ		t͡ɕ			
					tsʰ				ʈʂʰ		t͡ɕʰ			
sičnik prip.					s	z	ʃ	ʒ		ʐ				
					s				ʂ		ɕ			
nesič. prip.			f										x	
			f										x	
drsnik				ʋ								j		
					ɹ					ɻ				
enojni vibr.						r								
stran. drsnik						l								
						l								

Legenda:

NN: nezveneči nezvočniki; ZN: zveneči nezvočniki

zelena polja: slovenski fonemi

modra polja: kitajski fonemi

brezbarvna polja: relevantne alofonske uresničitve slovenskih fonemov

oranžna obroba: sporne vrednosti kitajskih fonemov

Da bi dosegli sistematično pretvorbo med soglasniki, predlagamo naslednje ustreznice, ki so prikazane v Tabeli 2.

Tabela 2: Povezave med slovenskimi in kitajskimi soglasniki (fonemi)

mesto artik.	izdišni soglasniki												
	ustnični				koronalni				nebni				grlni
	dvo-ustnični		zobno-ustnični		dlesnični		zadlesnični		retrofleksni		nebni		mehko-nebni
način artik.	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN
nosnik		m / m				n / n								ŋ / ŋ
zapornik	p / p / pʰ	b			t / t / tʰ	d							k / k / kʰ	g
sičnik zlitnik					ts / ts / tsʰ	dz	t͡ʃ	d͡ʒ	t͡ʂ / t͡ʂʰ	t͡ɕ / t͡ɕʰ				
sičnik prip.					s / s	z	ʃ	ʒ	ʂ	ʐ / ɕ				
nesič. prip.			f / f										x / x	
drsnik			ʋ	w	ɻ				ɻ		j	ɥ		
enojni vibr.					ɾ									
stran. drsnik					l / l									

Legenda:

NN: nezveneči nezvočniki; ZN: zveneči nezvočniki

zelena polja: slovenski fonemi

modra polja: kitajski fonemi

brezbarvna polja: relevantne alofonske uresničitve slovenskih fonemov

oranžna obroba: sporne vrednosti kitajskih fonemov

Skoraj vsi soglasniki dobijo svoje enovite ustreznice, izjema so naslednji pari soglasnikov, ki jim pripišemo en sam slovenski soglasnik:

- kitajskima [t͡ɕ] in [t͡ʂ] ustreza slovenski [d͡ʒ]
- kitajskima [t͡ɕʰ] in [t͡ʂʰ] ustreza slovenski [t͡ʃ]
- kitajskima [ɕ] in [ʂ] ustreza slovenski [ʃ]

Za lažjo razumljivost te rešitve poenostavimo fonetično-fonemski zapis v fonemski zapis (Tabela 3).

Tabela 3: Povezave med slovenskimi in kitajskimi soglasniki (grafemi)

mesto artik.	izdišni soglasniki													
	ustnični				koronalni				nebni				grlni	
	dvo-ustnični		zobno-ustnični		dlesnični		zadlesnični		retrofleksni		nebni		mehko-nebni	
način artik.	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN	NN	ZN
nosnik		m / m				n / n								ng
zapornik	p / b / p	b			t / d / t	d							k / g / k	g
sičnik zlitnik					c / z / c	dz	č	dž	zh	ch	j	q		
sičnik prip.					s / s	z	š	ž	sh	r		x		
nesič. prip.			f / f										h / h	
drsnik				v / w		r				ɻ	j			y
enojni vibr.						ɾ								
stran. drsnik						l / l								

Legenda:

NN: nezveneči nezvočniki; ZN: zveneči nezvočniki

zelena polja: slovenski fonemi

modra polja: kitajski fonemi

brezbarvna polja: relevantne alofonske uresničitve slovenskih fonemov

oranžna obroba: sporne vrednosti kitajskih fonemov

Do odločitve, da pare *ch-q, sh-x* in *zh-j* združimo v isti soglasnik, smo prišli po proučitvi teh glasov v vseh zlogih standardne kitajščine. Analiza je pokazala, da so ti soglasniki v popolni komplementarni distribuciji. Na primer, pred končajem *-a* lahko stojijo le *zh/ch/sh*, pred končajem *-ia*[7] pa samo *j/q/x*. Slovenska izgovarjava [ča] ustreza kitajskemu *cha*, slovenska izgovarjava [čja] pa kitajskemu *qia*.

V prid tej rešitvi, ki ne sledi zgolj fonetični podobi jezika, temveč upošteva tudi grafično reprezentacijo, govori tudi minimalno odstopanje med črkami. Nesmislelno bi bilo stremeti k identični glasovni podobi in kitajskemu 'b' pripisati slovenski 'p',

---

[7] V kitajščini se zapis *-ia* bere kot drsnik in samoglasnik [ja].

kitajskemu 'p' pa dodajati črko h za doseganje aspiracije 'ph'. S tem bi se obenem že nekoliko oddaljili od željenega [pʰ]. Če bi že na vsak način želeli dodati informacijo o pridihnjenosti, bi lahlo uporabili sočno prispodobo, ki jo ponuja Golob.

> *»Da bʹte koklja brcnila!«* Hvalabogu za zapornike, saj so odlični blažilci stresa, pa tudi izrazne fleksibilnosti jim ne manjka! Moč čustvene napetosti tako lahko izrazimo z bojem med iztiskanjem zraka in tesno zaprtimi ustnicami, za katerimi lahko pritisk naraste tudi do meje, ko obraz pordeči. Takrat nekoliko popustimo togost ustnic, da lahko glasove obogatimo z veliko mero piša (aspiracije) in učinek je zagotovljen: [ˌphɔrkhaˈdhuʃ].

Na tem mestu omenimo še eno podrobnost glede zapisa v oglatih oklepajih. Čeprav se načeloma med oglate oklepaje piše *fonetični* in ne *fonemski* zapis, v slovenski razlagi izgovarjave uporabljamo slednjega, pri čemer upoštevamo tudi zgoraj omenjena razmerja med glasom in črko.

Spomnimo se, da črka 'g' zapisuje glas 'k' pred nezvenečim nezvočnikom. Po tem principu švedsko besedo 'án**g**strem' oz. 'ân**g**ström' zapišemo [án**k**strem]. Podobno tudi angleško besedo 'lon**g**play' zapišemo [lôn**k**plêj].

Pričakovali bi, da bodo v skladu s § 1069 tudi črke 'g' na koncu besede zapisane kot [k], vendar novejše smernice uporabljajo mnogo boljšo rešitev:

benchmarkin**g** [bênčmárkin**g-**]
bôdibíldin**g** in bodybuildin**g** [bôdibíldin**g-**]
bowlin**g** [bôu̯lin**g-**] (FRAN)

Iz tega zapisa je jasno razvidno, da pri pregibanju ohranjamo [g], obenem pa vemo, da se v imenovalniku zadnji glasovi teh besed preberejo kot [k]. Razumevanje takega zapisa razreši vse dileme, kako zapisovati kitajske zloge, ki se končajo na -ng. Priimek 'Chang' tako zapišemo [čáng-] in ne [čánk]. Razumljivo postane tudi, zakaj ime kitajkega astronoma Wan**g** Chon**g**a zapišemo kot [ván**k** čún**g-**] in ne kako drugače.

Poznavanje prilikovanjskih (asimilacijskih) procesov je koristno ne le znotraj lastnega jezika, temveč tudi pri učenju in poučevanju tujih jezikov. Na vsaj srednji zahtevnostni stopnji bi se morali učenci/študenti zavedati, da prihaja v besedah *shénme* ali *guǎngbō* do sprememb izgovarjave. Za celovitejše razumevanje te tematike bi bilo koristno izdelati primerjalno študijo glasovnih procesov obeh jezikov.

4   Zaključek

Po številu soglasnikov sta slovenščina in kitajščina primerljiva jezika, ki se razlikujeta predvsem po pomensko-razločevalni funkciji zvenečnosti in pridihnjenosti. Predlagane rešitve iz Tabele 3 strnimo v krajše navodilo, primerno za slovenskega bralca, ki skuša prebrati zapis soglasnikov v pinjinu v okviru slovenskih knjižnih glasov.

1. Črke *b, c, d, f, g, h, k, l, m, n, p, s* in *t* izgovarjamo tako kot v slovenščini.

2. Razlika med zapisom v pinjinu in izgovorom v slovenščini:

   pinjin   >   izgovor
   r   [ž]
   ch, *q*   [č]
   sh, *x*   [š]
   zh, *j*   [dž]
   z   [dz]
   w   [v]
   y   [j]
   ng   [ng-]

Poudariti moramo, da zgolj taka oblika zapisa navodil za izgovarjavo soglasnikov še ni dovolj, saj potrebujemo pregled vseh končajev in njihove kombinacije s soglasniki v vzglasju. Zaradi večplastnosti in obsežnosti problematike ta del raziskave v okviru Pravopisne komisije pri SAZU in ZRC SAZU še ni zaključen.

Sofinanciranje raziskave

Reference

Chinese pronunciation wiki (2015-2018). IPA. Dosegljivo na
https://resources.allsetlearning.com/chinese/pronunciation/IPA

Comparison of English and Mandarin (Segmentals). (2019). A Corpus-based English Pronunciation Learning Website. The Education University of Hong Kong. Dosegljivo na http://corpus.eduhk.hk/English_Pronunciation/?page_id=328

Defense Language Institute (1974). *A Contrastive Study of English and Mandarin Chinese*. Washington, D.C.: ERIC Clearinghouse.

Deroń, M. (2018). *IPA 2018 i-charts*. Dosegljivo na
    https://linguistics.ucla.edu/people/keating/IPA/inter_chart_2018/IPA_2018.html

Dowse, J. (2018). *IPA Charts with Audio*. Dosegljivo na https://jbdowse.com/ipa/

EasyPronunciation.com (2013-2019). Chinese Pronunciation Tool - Pinyin Translator.
    Dosegljivo na https://easypronunciation.com/en/chinese-pinyin-phonetic-transcription-
    converter#phonetic_transcription

Golob, N. (neobjavljeno). Glasovna podoba japonskega jezika.

Golob, N., & Petrovčič, M. (2018). Hokkaido Pumpkins and Huawei Phones: Anti-hiatus
    Tendencies in Slovene. *Acta Linguistica Asiatica*, *8*(2), 63-82.
    https://doi.org/10.4312/ala.8.2.63-82

Huang, K. (2018). Phonological Identity of the Neutral-tone Syllables in Taiwan Mandarin: An
    Acoustic Study. *Acta Linguistica Asiatica*, *8*(2), 9-50. https://doi.org/10.4312/ala.8.2.9-50

International Phonetic Association, (Ur.) (1999). *Handbook of the International Phonetic
    Association*. Cambridge: Cambridge University Press.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*.
    Berkeley: Univ. of California Press.

Lingvo.info (2019) *Fonetika, inventar zvokov. Zvoki, ki jih govorimo*. Dosegljivo na
    http://lingvo.info/sl/babylon/phonetics

Sailaja, P. (1999) Syllable Structure of Telugu. *14th International Congress of Phonetic Sciences
    (ICPhS-14)*, 743-746. Dosegljivo na
    https://www.internationalphoneticassociation.org/icphs-
    proceedings/ICPhS1999/papers/p14_0743.pdf

Švarný, O., & Uher, D. (2001). *Hovorová čínština: úvod do studia hovorové čínštiny*. 2. Olomouc:
    Univerzita Palackého.

Toporišič, J. (ur.) (2001). *Slovenski pravopis*. Ljubljana: Slovenska akademija znanosti in
    umetnosti in Znanstvenoraziskovalni center SAZU, Inštitut za slovenski jezik Frana
    Ramovša. Dosegljivo na http://bos.zrc-sazu.si/c/sp/sp2001_pravila.pdf

Uher, D., Jin, X., & Slaměníková, T. (2016). *Učebnice čínské konverzace II*. Praha: Leda.