Institut für Sprache und Information

The Influence of Orthography on Spoken Word Processing

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie (Dr. phil.)

durch die Philosophische Fakultät der

Heinrich-Heine-Universität Düsseldorf

Vorgelegt von

LEILA NAYERNIA

Aus

SHIRAZ, IRAN

Betreuer:

1. Prof. Dr. Dr. Peter Indefrey
2. Prof. Dr. Ruben van de Vijver

Düsseldorf, September 2022

D61

“**Brush up your English**

…

Beware of heard, a dreadful word

that looks like beard and sounds like bird,

and dead – it’s said like bed, not bead.

For goodness sake, do not call it deed!”

…

T.S. Watt (1954)

Abstract

The influence of orthography on speech processing was investigated in five experiments. At the perceptual level, (pseudo)words containing phonological units with or without orthographic realizations were successfully discriminated by Persian literates and illiterates in AX and ABX discrimination experiments, suggesting no orthographic effects at this level. At the lexical level, orthographic effects were observed in a German AX discrimination experiment and a Persian primed lexical decision experiment. German homographic words were discriminated less accurately than German non-homographic words by German literates, but homography had no influence on the discrimination of German pseudowords suggesting an orthographic effect at the lexical level. Spelling similarity facilitated lexical decision latencies for embedded target words when they were preceded by embedding words, however, no facilitation was observed when embedding and embedded words had dissimilar spellings. At the intentional level, phonological units of words were manipulated more accurately when they were spelled than when they were not in Persian and German phoneme reversal experiments. In a phoneme monitoring experiment on Persian, phonemes with orthographic realizations were detected more accurately than phonemes without orthographic realizations when the phonemes were of mid or short durations, but phoneme detection was not a function of orthographic realizations in phonemes with long duration, suggesting that phonological and orthographic features of spoken words influence spoken word processing at the lexical and the intentional level.

Spelling transparency effects reported from written word naming experiments (Rahbari, 2019; Bakhtiar & Weekes, 2015) were not observed in a semantic categorization experiment using Persian written words, suggesting that phonological inconsistencies of graphemes can impair naming but not the semantic processing of written words. In another experiment, written pseudowords were named more accurately when they contained long vowels than when they contained pointed short vowels, suggesting reduced phonemic awareness for phonological units (Persian short vowels) that are not conventionally represented in spelling.

Thus, phonological information can influence written word processing and spelling information can influence spoken word processing. A consistent pronunciation for the same spelling unit and a consistent spelling for the same phonological unit facilitate word processing. In sum, the findings suggest that lexical entries may undergo a developmental restructuring in terms of the combined phonological and orthographic features of words leading to a phonographic structure of the lexicon.

Zusammenfassung

Der Einfluss der Orthographie auf die Sprachverarbeitung wurde in fünf Experimenten untersucht. Auf perzeptueller Ebene wurden (Pseudo)wörter, die Segmente mit oder ohne orthographische Realisierung enthielten von persischen Hörern in AX- und ABX-Diskriminationsexperimenten unabhängig von der Lesefähigkeit der Hörer erfolgreich unterschieden, sodass auf dieser Ebene keine orthographischen Effekte vorzuliegen scheinen. Auf der lexikalischen Ebene wurden orthographische Effekte in einem deutschen AX-Diskriminationsexperiment und einem persischen lexikalischen Entscheidungsexperiment mit Primewörtern gefunden. Deutsche Hörer unterschieden homographe Wörter weniger korrekt als nicht-homographe Wörter. Im Gegensatz dazu, hatte Homographie keinen Einfluss auf die Unterscheidungsfähigkeit für deutsche Pseudowörter. Im persischen lexikalischen Entscheidungsexperiment verkürzte eine ähnliche Schreibweise die Entscheidungslatenz für Wörter, die in vorangehende Primewörter eingebettet waren. Diese Beschleunigung wurde bei unähnlicher Schreibweise von Wörtern und Primewörtern nicht gefunden. Auf der intentionalen Ebene wurden in persischen und deutschen Wortumkehrungsexperimenten phonologische Einheiten von Wörtern korrekter umgeordnet, wenn sie orthographisch realisiert waren, als wenn das nicht der Fall war. In einem persischen Phonemerkennungsexperiment wurden Phoneme mit orthographischer Realisierung korrekter erkannt als solche ohne orthographische Realisierung, wenn die Phoneme von mittlerer oder kurzer akustischer Dauer waren. Bei Phonemen mit längerer akustischer Dauer spielte die orthographische Realisierung dagegen keine Rolle für die Erkennung. Insgesamt legen die Ergebnisse nahe, dass orthographische sowohl als phonologische Eigenschaften die Verarbeitung von gesprochenen Wörtern auf lexikalischer und intentionaler Ebene beeinflussen.

Bekannte Effekte der orthographischen Transparenz beim Vorlesen von Wörtern (Rahbari, 2019; Bakhtiar & Weekes, 2015) wurden in einem Experiment zur semantischen Kategorisierung von geschriebenen persischen Wörtern nicht beobachtet, sodass phonologisch inkonsistente Grapheme zwar die Benennung nicht aber die semantische Verarbeitung von geschriebenen Wörtern nachteilig zu beeinflussen scheinen. In einem weiteren Experiment wurden persische Pseudowörter korrekter

vorgelesen, wenn sie lange Vokale enthielten, als wenn sie diakritisch markierte kurze Vokale enthielten. Dieses Ergebnis legt eine geringere phonematische Bewusstheit für phonologische Einheiten (persische kurze Vokale) nahe, die in der konventionellen Schreibweise nicht orthographisch repräsentiert sind.

Insgesamt fanden sich also sowohl Einflüsse phonologischer Information auf die Verarbeitung geschriebener als auch Einflüsse der Schreibweise auf die Verarbeitung gesprochener Wörter. Grapheme mit konsistenter Aussprache und Phoneme mit konsistenter Schreibweise erleichtern die Verarbeitung von Wörtern. Die Zusammenschau aller Ergebnisse dieser Dissertation legt nahe, dass im Laufe des Erwerbs der Lese- und Schreibfähigkeit die Lexikoneinträge durch die Kombination phonologischer und orthographischer Eigenschaften so umstrukturiert werden, dass eine phonographische Lexikonstruktur entsteht.

Acknowledgement

The completion of this thesis was not possible without the assistance of several special people. Hence, I would like to take this opportunity to thank them.

I am profoundly indebted to my supervisor, Prof. Dr. Dr. Peter Indefrey, for his valuable and constructive advice, helps, and meticulous editing without which the completion of this research project was impossible.

An immense thank you to my second supervisor, Prof. Dr. Ruben van de Vijver, who encouraged me in finalizing this dissertation.

I would like to show my gratitude to my defense committee who accepted my invitation to be here, Prof. Dr. Ingo Plag and Prof. Dr. Kilu von Prince, without whom I could not have undertaken this journey.

I would also like to thank the ex-manager of the language lab at HHU, Dr. Alexandra Redmann, who helped me in operationalizing the experiments with German Participants. My thanks go further to Tim Marton, the secretary of the linguistic department at HHU, who was responsive to my questions and Emails.

It would be remiss of me not to express my gratitude to my family: my late father who inspired me 25 years ago, before his death, to never stop education; Mohsen who was with me in all steps of running the experiments; my mother and my sister who offered their encouragements through phone calls and did not let me quit in ups and downs.

I wish to extend my appreciation to Dr. Vondran and his wife who offered me the hospitality of their home during the first years of my study.

You are all amazing!

Table of contents

List of Figures

List of Tables

List of abbreviations

Abbreviation	Meaning
Abbr.	Abbreviation
Avg.	Average
BIAM	Bimodal interactive activation model
C	Consonant
DRC	Dual Route Cascaded model
Cons.	Consistent
Ex.	Example
Exp.	Experiment
Fig.	Figure
Freq.	Frequency
H	Homograph
h	Hour
HF	High frequency
Incons.	Inconsistent
L	Long vowel
LF	Low frequency
M	Mean
min	Minute
M.P.	Minimal pair
ms	Millisecond
No.	Number
P	Pseudoword
p.	pointed
R	Real word
S	Short vowel
SD	Standard deviation
u.	unpointed
V	Vowel
W	Word

Chapter 1: Introduction

1-1. Word processing

One of the hot debates in the domain of language processing is how different types of word knowledge such as phonological knowledge and orthographic knowledge contribute to the process of word recognition[1]. Models differ in how various types of knowledge interact and how the information flows between the levels of processing. In autonomous or feedforward models, it is assumed that the information flows in one direction from the lower levels (extracting features from input) to the higher levels while processing words ((Indefrey & Levelt, 2004), p.109; (Norris, 1999)). In interactive models, however, the information flows in both directions, that is, information from higher levels can be feed-backed to lower levels (McClelland, 1987).

Some models of written word recognition maintain that written words can be recognized without any phonological mediation (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Rapp, Benzing, & Caramazza, 1997). Others propose that phonological knowledge is obligatorily activated in the process of written word recognition (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005; Frost, 1998; Van Orden, 1990). As for the information flow between the levels of processing, some models for written word recognition allow for information to flow from bottom to top (Indefrey & Levelt, 2004; Frost, 1998), and others allow for information flow in both directions between the lexical and prelexical levels (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001).

Some models of spoken word recognition maintain that spoken words activate orthographic knowledge automatically (Diependaele, Ziegler, & Grainger, 2010; Ziegler, Petrova, & Ferrand, 2008). Others assign no explicit role to the orthographic knowledge in spoken word processing (McQueen, Cutler, & Norris, 2000; Norris, 1999; McClelland & Elman, 1986; Marslen-Wilson & Welsh, 1978). The latter group can be further subdivided such that some completely rule out the interaction of phonological

[1] "The process of going from the input string to the selection of a single item stored in the lexical memory" (Lupker, 2008), p. 39.

and orthographic knowledge in spoken word processing, and others do not. The TRACE (McClelland & Elman, 1986) and Cohort (Marslen-Wilson & Welsh, 1978) models completely rule out any orthographic effect in the process of spoken word recognition. The Merge model (Norris, McQueen, & Cutler, 2000; Norris, 1999) considers a strategic role of orthographic knowledge at the postlexical level when the task involves a decision mechanism and when orthographic knowledge can help task performance (Norris, McQueen, & Cutler, 2000). As for the information flow between the levels of processing, some models of spoken word recognition only allow a unidirectional flow of information between the levels of processing from lower to higher levels (Indefrey & Levelt, 2004; Norris, McQueen, & Cutler, 2000; Norris, 1999). Others let the information flow in both directions (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005; McClelland & Elman, 1986). Hence there is no consensus whether the contribution of orthographic knowledge is automatic[2] /strategic or (post)lexical/ sublexical.

The findings of some experimental studies support the role of orthographic knowledge in spoken word processing (e.g., (Pattamadilok, Morais, & Kolinsky, 2011; Ziegler & Ferrand, 1998)); thus supporting models that allow for an interaction between phonological and orthographic knowledge during spoken word processing, such as the Bimodal Interactive Activation Model, i.e., BIAM (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005). On the other hand, there are also experimental studies suggesting that the influence of orthography may be limited to lexical and post lexical processing levels, when experimental tasks activate orthographic knowledge in addition to phonological knowledge (Cutler & Davis, 2012; Cutler, Treiman, & Van Ooijen, 2010). These studies support the Merge model for spoken word recognition (Norris, McQueen, & Cutler, 2000; Norris, 1999). A detailed description of the experimental evidence is presented in the Literature review Chapter.

The advocates of the BIAM (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005) believe that orthographic knowledge and

[2] Automatic effects or processes are unconscious, unintentional, fast, and insensitive to interference from other ongoing mental activity/activities (Posner & Snyder, 1975).

phonological knowledge are tightly bound together. Ehri (1985) makes a close analogy between music and speech to illustrate the "amalgamation" of the phonological and orthographic knowledge that is established during learning to read and write. By learning to read sheet music, one learns to visualize what is sung and played by instruments; and by learning to read/write, one learns a means of visualizing what is heard and said. In both cases, what is invisible and has a temporal feature (music/speech) is represented by a visual-spatial form (Ehri, 1985). Thus by learning to read/write, a visual representational system is bound to speech.

The links between phonological representations and the corresponding orthographic representations (symbolic visual representations) are established and consolidated during the reading/writing process (Ehri, 2005). This phenomenon can lead to a form of unitization; so that when one is encountered, the other is automatically co-activated (Diependaele, Ziegler, & Grainger, 2010; Ziegler, Petrova, & Ferrand, 2008; Ziegler, Ferrand, & Montant, 2004; Ehri, 1985). In other words, phonological representations activate the corresponding orthographic representations and vice versa. In such an interrelated system, it can be considered a processing advantage, when the link between phonology and orthography is one-to-one, that is when a given phonological unit is represented always by the same orthographic unit. For example, the phoneme /m/ is represented by <m> in English. However in case of any mismatch _ the phoneme /j/ has no exact orthographic representation in 'fuse'_ (Van Ooijen, Cutler, & Norris, 1992; Ehri, 1985) or inconsistency _ the phoneme /s/ is inconsistently represented in 'cycle' and 'system'_ (Ziegler & Ferrand, 1998) in the links between phonology and orthography, processing can be hindered. The automatic activation of multiple orthographic representations for the same phonological unit may lead to a competition that slows down speech processing (Ziegler, Ferrand, & Montant, 2004).

Alternatively, it could be the case that phonological representations that have no orthographic representations or inconsistent orthographic representations are different in quality or nature from phonological representations that are consistently spelled. The phonological representations for these two groups of sounds might be of a different nature and quality, because phonological representations may be restructured in terms of orthographic knowledge (Taft, 2011; Goswami, 2000) in a way that phonological representations of the phonemes with the exact or consistent orthographic representation are fully specified. For example, Taft (2011) shows that the word ‘swap’ in English may develop multiple phonological representations /swɔp/ and /swæp/ after literacy. Such multiplicity in phonological code may affect the phonological processing of that word as compared to words, which develop a unique phonological representation after literacy (such as ‘stop’ /stɔp/). Thus, having multiple/no orthographic representations can influence the nature of the phonological representations themselves. Hence, consistent or exact orthographic representations can improve the precision of the phoneme boundaries (Kolinsky, 2015; Hoonhorst, et al., 2011) or the phonological processing (Silva, Faisca, Ingvar, Petersson, & Reis, 2012). In languages such as English (Ziegler, Petrova, & Ferrand, 2008), French (Pattamadilok, Morais, De Vylder, Ventura, & Kolinsky, 2009), and Portuguese (Ventura, Kolinsky, Pattamadilok, & Morais, 2008), it is well documented that an inconsistent relationship between phonology and orthography (one phonological unit is represented by multiple orthographic units) can hinder speech processing as compared to a consistent relationship. However, effects of an incomplete relationship between phonology and orthography on speech processing have received less attention. This might be because in the languages that are mainly under investigation (French, English, and Portuguese), almost all sounds are orthographically represented. One exception is a study by Van Ooijen and colleagues (Van Ooijen, Cutler, & Norris, 1992) who found that when a sound had no orthographic realization (e.g., /j/ in ‘fuse’), it was detected less accurately than when the same sound did have an orthographic realization, e.g., /j/ in ‘yell’.

The aim of the present thesis is to provide new empirical evidence for how incomplete and inconsistent relationships between phonology and orthography may influence speech processing. This question was investigated by conducting experiments in Persian. Persian uses a non-Roman script with a highly opaque relationship between phonology and orthography. This opacity provided the opportunity to study effects of orthography on speech processing without orthographic manipulations becoming salient in auditory tasks.

There are several ways by which one can find whether orthography can modulate speech processing. Firstly, by comparing the performance of literates when the link between phonology and orthography is consistent, inconsistent, or incomplete, it can be shown whether orthographic knowledge affects speech processing or modulates the phonological system. Additionally, by comparing the performance of literate and illiterate populations while processing speech, it can be found how introducing a visual representational system (orthographic representations) can influence the processing, nature, and organization of phonological information.

The mapping or link between phonology and orthography can be of three types: consistent, inconsistent, or incomplete. In the consistent mapping, there is a one-to-one relationship between a phonological unit and its corresponding orthographic unit (e.g., /m/ and /ʌk/ are represented consistently by <m> and <-uck> respectively in English). The inconsistent relationship refers to the cases where a phonological unit is represented by multiple orthographic units[3]. The phonological units /k/ and /-i:d/, for example, are orthographically represented by <k, c, q> and <-eed, -ead> respectively in English. When a phonological unit is not represented by any orthographic unit, the relationship between phonology and orthography is incomplete (e.g., the sound /j/ has no orthographic realization in <fuse> /fju:z/).

According to the Bimodal Interactive Activation Model - BIAM, (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005), orthographic knowledge affects spoken word processing automatically, moreover, the influence of orthography is expected to be observed at both lexical and sub-lexical levels due to the

[3] Since the effect of orthography on spoken word recognition is assumed to emerge through a feedback mechanism from the orthographic system to the phonological system, Ziegler and colleagues (Ziegler, Petrova, & Ferrand, 2008; Ziegler & Ferrand, 1998) consider inconsistencies as cases where one phonological unit is represented by multiple spellings (sound-to-spelling inconsistency). The consistency effect occurs in this sense by competition between/among multiple spellings. Cases where multiple phonological units are represented by the same orthographic representation (spelling-to-sound inconsistency) are excluded from the orthographic consistency effect by Ziegler et al. In fact, they believe that the sound-to-spelling and spelling-to-sound inconsistencies would selectively influence the spoken and written word recognition respectively. By contrast, the Recurrent Feedforward and Feedback Model (Stone, Vanhoy, & Van Orden, 1997) predicts that any inconsistency in spelling-pronunciation mapping would lead to slower reaction times in speech processing. In an interactive language system, in which the information flows in both directions between phonology and orthography, both types of inconsistencies are expected to influence speech processing. In the spelling-to-sound inconsistency, the consistency issue may emerge through restructuring phonological representations or through a feedback mechanism from orthography to phonology which reduces the activation level of the multiple sounds represented by the same orthographic representation as compared to the consistent cases (Taft, Castles, Davis, Lazendic, & Ngye-Hoan, 2008).

existence of cross links between phonology and orthography at both levels and due to the feedback from the lexical level to the sub-lexical level. On the other hand, other models of spoken word recognition consider no role or a strategic role for orthographic knowledge in spoken word processing. In the latter case (strategic role for orthography), the influence of orthographic knowledge is limited to the post lexical level whenever a task involves a decision process and induces the application of orthographic knowledge. The present study investigates how a mismatch or inconsistency of the links between phonology and orthography influences Persian speakers' speech processing, and how their performance can be interpreted based on models of spoken word recognition. The relevant models are, therefore, presented briefly in the following section. Since links between phonology and orthography are supposed to emerge during learning to read and write, Section 1-3 will briefly present current models of written word recognition.

1-2. Models of spoken word recognition

As discussed in the previous section, models of speech processing differ with respect to possible influences of orthographic knowledge on the speech processing. The most influential models are briefly presented in this section.

1-2-1. Cohort model

In the Cohort model (Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978), all lexical entries that share the same initial phonological segment(s) with a spoken input are activated. These lexical entries form the so-called 'cohort' of word candidates. For example, by hearing /d…/, all lexical entries starting with /d/ are activated, e.g., 'door, deal, dean, duck, etc.' When the second phonological segment of the spoken input arrives, some of the activated lexical entries that do not match the spoken input anymore are removed from the cohort. For example, by hearing /di…/, 'door' and 'duck' are deactivated, because they do not match to the spoken input anymore. This process continues until just one lexical entry remains activated (uniqueness point). In this model, cohorts are constituted solely by the phonological information of the spoken input. Thus, this model assumes no interaction between the phonological and orthographic knowledge in the process of spoken word recognition.

1-2-2. TRACE model

The TRACE model (McClelland & Elman, 1986) considers three layers of processing with continuous uptake of information across layers: a feature layer, a phoneme layer, and a word layer. The activated nodes at each layer activate the compatible nodes at the higher level (excitatory links displayed by arrows in Figure 1). Features in the input activate the corresponding phonemes, and the phonemes activate the related word candidates. The words that are matched in sound to the spoken input in onset or rhyme are activated (in contrast to the Cohort model (Marslen-Wilson & Welsh, 1978) in which only those words are activated that match to the input in initial sounds, see above). Thus, the input 'deal' /di:l/ activates 'door' /dɔ:r/ as well as 'feel' /fi:l/ among others. There are inhibitory links between nodes within each layer (displayed with lines ended in circles).

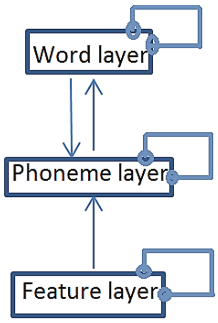


Figure 1. Simple schematic view of the TRACE model for spoken word recognition (McClelland & Elman, 1986).

In the TRACE model, the information in the phoneme layer activates nodes at the word layer (feedforward connection) and the activated word candidates send feedback to the phoneme layer. Thus, the information at the higher lexical layer can influence the sublexical perception by sending online feedback. Hence, it is considered as an interactive model for spoken word recognition (McClelland, Mirman, & Holt, 2006). The TRACE model, like the Cohort model, does not consider any explicit role for orthographic knowledge during spoken word recognition.

1-2-3. Merge model

In the Merge model (McQueen, Cutler, & Norris, 2000; Norris, 1999), there are three basic units (levels) dealing with sublexical (input), lexical, and decision processing. The activation of information in the sublexical/ prelexical[4] unit (labeled 'input' in Figure 2) spreads into the corresponding nodes at the lexical level. Then, the activated nodes at the lexical level spread into the corresponding nodes at the decision level. Likewise, the activation of information in the input can directly spread into the corresponding nodes at the decision level.

There are excitatory links (unidirectional solid lines) between the units and inhibitory links (dotted lines) among the nodes within lexical and decisional units. The lack of any inhibitory link at the prelexical level serves to avoid any categorical decision at this level. In this way, the genuine information is passed from sublexical nodes to the higher-level nodes as it is received, and no information is lost or overridden. Thus, all sorts of information in the input e.g., ambiguities, are preserved in the information flow, and passed on to the lexical and decision levels.

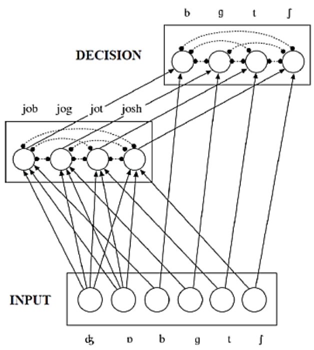


Figure 2. Merge Model for spoken word recognition (McQueen, Cutler, & Norris, 2000), p.47.

The Decision unit continuously receives information from lexical as well as prelexical units and merges them to arrive at a decision which is biased depending on the task demands. That is, different weights are given to information arriving from the lexical and prelexical levels based on the task demands. For example, lexical knowledge might be weighted more strongly in lexical decision tasks. When prelexical information is phonemically ambiguous, lexical information may serve as the basis (receive more weight) for decision-making. For instance, an ambiguous 't-d' phonemic (phonetic)

[4] Considering the terms 'prelexical' and 'sublexical', the former refers to a processing level and the latter to a processing unit. These two terms might be used interchangeably in this dissertation since at the prelexical level, sublexical units are processed.

input may be ‘considered’ or ‘interpreted’ (rather than ‘perceived’[5]) as ‘d’ in the context ‘-ice’ by the decision mechanism. The decision is less perceptually and more lexically biased in this case. It is lexically biased since ‘tice’ would yield a nonword, whereas ‘dice’ is a real word. The decision mechanism processes information explicitly; that is, any decision, e.g., phoneme detection, is made explicitly considering all sorts of available resources.

The Merge model is different from other models of spoken word recognition in that it incorporates a decision-making mechanism. The decision mechanism takes all available information (prelexical, lexical, and contextual information) into account; but assigns the different kinds of information different weights based on the task demand. According to Norris and colleagues (Norris, McQueen, & Cutler, 2000), the Merge model is the most parsimonious and at the same time a natural model which can accommodate data from speech recognition to phoneme decision tasks[6] that are not part of the normal speech recognition process. Based on this model, phonemic decisions are made at the decision level. Speech perception and phonemic decisions have different requirements and cannot be performed based on the same units. That is why illiterates have problem in making phonemic decisions while they perceive speech without problem. At the decisional level, orthographic knowledge can play a role in making phonemic decisions.

1-2-4. Bimodal Interactive Activation Model (BIAM)

In BIAM (Grainger, Muneaux, Farioli, & Ziegler, 2005), there are two levels: sublexical and lexical (the rows in Figure 3) in the auditory and visual modalities (the columns in Figure 3). At the lexical level (word), there is one direct bidirectional link between orthography and phonology, and another bidirectional link exists between orthography and phonology at the prelexical level via the O<>P unit (representing grapheme-

---

[5] In the TRACE model, the input ambiguities are ‘perceived’ based on the lexical knowledge due to the feedback from the lexical to the prelexical level.

[6] Phoneme decision tasks refer to tasks in which participants should decide whether an input includes a target phoneme. They may also be asked to categorize spoken inputs as a phoneme.

phoneme correspondences). Thus, according to BIAM, orthographic effects are expected at both levels of speech processing. In the visual modality, orthographic units feed into the orthographic lexicon; while in the auditory modality, phonological units feed into the phonological lexicon.

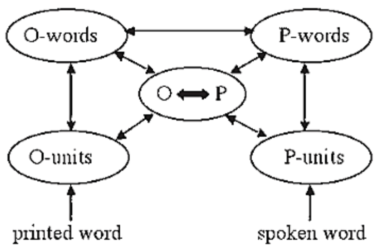


Figure 3. Architecture of the Bimodal Interactive Activation Model (BIAM) for word recognition (Grainger, Muneaux, Farioli, & Ziegler, 2005), p. 982. The O<>P central unit is for orthography-to-phonology and the reverse conversion.

In BIAM, the information flows from the sublexical to the lexical level and back within modality. Thus, lexical information can constrain the sublexical perception.

1-3. Models of written word recognition/ reading

Reading is a cognitive process that entails visual (printed) word identification or recognition (Perfetti, Van Dyke, & Hart, 2001). According to traditional models of printed word recognition, printed words are processed and mapped onto a word entry in the lexicon (Coltheart, 1978). This route can be phonologically mediated. Alternatively, phonological mediation may be bypassed on a direct route (see Figures 3 and 4).

One of the features that varies cross-linguistically is the transparency of the orthographic system which is defined in terms of the consistency of the grapheme-phoneme correspondences in languages with alphabetic writing systems (Perfetti, 1997). Whereas transparent orthographies are characterized by a *consistent* grapheme-phoneme correspondence, the consistency decreases in opaque orthographies. It has been shown that orthographic transparency can influence the reading strategies that readers of a language develop (Ziegler & Goswami, 2005; Seymour, Aro, & Erskine, 2003).

Frost (1998) believes that the prelexical processing of a printed word ends in generating an underspecified phonological code which varies in specifications based on the orthographic transparency. This sublexical phonological code is used for lexical access. In transparent languages, the sublexical phonological code is more detailed. Thus, it can be assumed that lexical access is faster and more efficient in transparent orthographies than in opaque orthographies. Moreover, readers in opaque orthographies depend on reading (coding) larger units, whereas in transparent orthographies the reading units are smaller in grain size (Ziegler & Goswami, 2005). In other words, in opaque orthographies multiple letters (graphemes) are converted into phonemic clusters, whereas in transparent orthographies the converted reading unit can shrink to single graphemes corresponding to single phonemes. According to Frost (1998), the size of the reading unit can also influence the reading efficiency and speed.

Other models assume that words are not represented in the lexicon; rather, they emerge as a result of parallel and distributed activation (Plaut, McClelland, Seidenberg, & Patterson, 1996) or the stabilization of dynamic patterns (Van Orden & Goldinger, 1994). According to the latter, patterns with phonological mediation can equilibrate faster than patterns without phonological mediation (direct grapheme-meaning mapping). Hence, fast written word identification is mediated by phonological information. The results of a reading experiment in Chinese, a language with a logographic writing system, show that even in this language phonological codes are generated while reading or recognizing printed words (Perfetti & Zhang, 1995), suggesting that the process in reading or written word recognition is highly dependent on the phonological representations of the spoken language (Perfetti, Zhang, & Berent, 1992).

The models described so far have been criticized because they are not sensitive to the internal structure of words (Perfetti, Van Dyke, & Hart, 2001). A two-cycle theoretical model presented by Berent and Perfetti (1995) postulates that the assembly is performed separately for vowels and consonants with a priority for consonant assembly[7]. This priority may be due to a higher consistency in the grapheme-phoneme correspondence for consonants as compared to vowels in English.

---

[7] According to this model, assembly of phonemes is not performed in a linear way, but assembly is performed based on some planes each of which deals with certain types of phonological information such

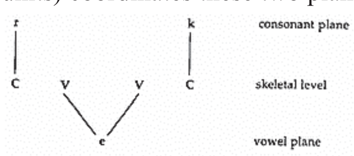
Thus, almost all models agree that the links between orthography and phonology are consolidated during reading/writing (e.g., (Stone, Vanhoy, & Van Orden, 1997; Van Orden, Jansen op de Haar, & Bosman, 1997; Van Orden & Goldinger, 1994). Data suggesting a fast activation of phonology in written word recognition have motivated the development of a model in which the visual input can directly activate the corresponding phonological information at sublexical and lexical levels, the Bimodal Interactive Activation Model for visual and spoken word recognition_ Figure 3 (BIAM; (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005).

In the Dual Route Cascaded (DRC) model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), three paths can be followed to read/recognize a written word (Figure 4). A lexical-semantic route (left path), a lexical-non-semantic route (direct path), and a non-lexical route (right path). The bilateral links between lexical components (phonological lexicon, orthographic lexicon, and semantic representation) indicate that information can reciprocate between them. Also, there are bidirectional links between letter representations and the orthographic lexicon as well as between phoneme representations and the phonological lexicon. Thus, the information flows forward and back between sublexical and lexical units.

In contrast to the BIAM (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005) which predicts the fast activation of phonology during written word recognition and reading, in the DRC model, the activation of phonological information can only occur fast via the lexical route (Diependaele, Ziegler, & Grainger, 2010). Through the non-lexical route (rule-based translation), the activation of lexical phonological information would occur late in the process of written word recognition and reading.

---

as features, syllable structures, and tones. Based on this model, assembly is performed in two cycles which assign vowels and consonants to two different planes. Then a skeleton (an abstract set of timing units) coordinates these two planes ((Berent & Perfetti, 1995), p.148):



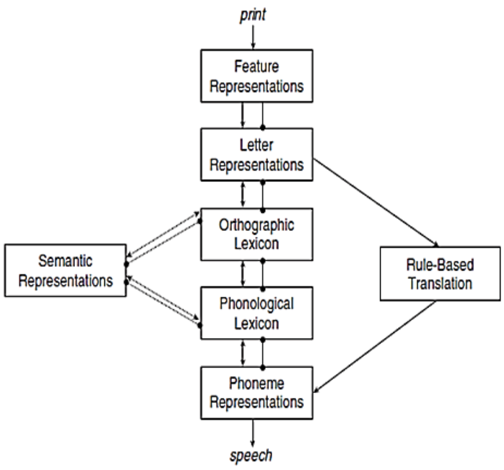


Figure 4. Dural Route Cascaded Model for written word reading and recognition (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), p. 213.

In cases where a fast activation of phonology is expected, the non-lexical route in the computational model may end up with incorrect pronunciations for exception or irregular words such as /heiv/ for 'have' (Diependaele, Ziegler, & Grainger, 2010). Diependaele and colleagues (Diependaele, Ziegler, & Grainger, 2010) criticize the DRC model for its weakness in explaining the fast activation of phonology in written word recognition, for which there is empirical evidence (e.g., (Rastle & Brysbaert, 2006; Lukatela & Turvey, 1994)). A printed word naming experiment with the masked phonological priming paradigm showed that 'bloo' primed 'BLUE' but 'blai' did not prime 'BLUE' (Lukatela & Turvey, 1994). That means 'blue' was named faster when primed by 'bloo'; however, the same facilitation was not observed in the orthographic control condition ('blue' primed by 'blai'). The early activation of phonological information was observed in both written word recognition and reading experiments (Rastle & Brysbaert, 2006).

In sum, phonology has an important role in the process of written word recognition, and it is activated fast during this process. The links created between orthography and phonology are assumed to consolidate after a long exposure to script which leads to the reciprocation of information between these two components of language. These considerations predict certain consequences for readers in opaque languages (such as Persian where the links between phonology and orthography are less consistent) after a long exposure to script.

First, their phonological lexicon may be redundant (Perfetti & Hart, 2002), that is, it includes two phonological codes, one derived from the spoken language and the other from the written language. Second, their phonological processing may be distorted by the simultaneous activation of orthographic information that is not always complete or consistent.

In all models of spoken and written word recognition discussed so far, the input is processed at several levels for recognition to take place. At each layer or level, different forms of auditory representations are processed. In the following section, I will present current models that make more detailed proposals for the kinds of representations to be distinguished for phonological processing. Each proposal has a unique characteristic which can help us understand and interpret the findings of the current study.

1-4. Levels of speech processing

Speech processing (alternatively may be called spoken word processing) takes place at different levels in which different forms of information are encoded and processed. Often, it is assumed that input is encoded and processed in subsequent steps, and each step has an output which serves as input for the next level (Indefrey & Levelt, 2004). Levels of processing need to be considered when dealing with language processing, because an effect (e.g., orthographic effect) may be observable only at a certain level and not necessarily across all levels of speech processing (e.g., (Chomsky & Halle, 1968)). Alternatively, orthographic information has been analogized to a virus which infects the whole language processing system (Frith, 1998). Thus, it is a point of disagreement in the literature whether orthographic information affects speech processing only at one or more specific levels, or whether the whole system of language processing is affected by orthographic information.

1-4-1. Processing levels in phonological processing

According to Taft and Hambly (1985), speech signals are decoded to be recognized. This decoding is performed at three different levels: an acoustic level, a phonetic level, and a phonological (phonemic) level. The amount of signal that is decoded at each time varies as well: single phone, phoneme, syllable, whole word, or sentence-like chunk (Mehler, Dommergues, Frauenfelder, & Segui, 1981). According to the so-called dual code hypothesis, prelexical processes are related to phonetic, and (post)lexical processes to phonological (phonemic) decoding. Phonetic codes that are derived from acoustic signals are abstract codes that do not reflect allophonic variations[8] (Foss & Blank, 1980). On the other hand, phonological codes are abstract codes stored in the lexicon. It should be mentioned that the "phonetic codes" presented by Foss and Blank are more abstract than the traditional description (see Footnote 8).

In the dual code model proposed by Foss and Blank (Foss & Blank, 1980), it is not clear what "*abstract* phonological codes" are (see (Taft & Hambly, 1985) for this critique): whether phonological codes are morphophonemic representations (Chomsky & Halle, 1968) or surface phonemic representations (Linell, 1979). According to Chomsky and Halle, morphologically related words share the same abstract phonemic representation, which is called morphophonemic representation, whereas the surface phonemic representation reflects the silent pronunciation of words.

About the effect of orthography on these representations, Chomsky and Halle believe that orthography is reflected in morphophonemic representations. As an example, the morphophonemic representation of MUSCLE contains a /k/ since this sound exists in the surface phonemic representation of MUSCULAR. Hence, no consensus has been reached about the nature of the phonological representations (Taft & Hambly, 1985).

In a series of experiments, Taft and Hambly (Taft & Hambly, 1985) showed that phonological representations are neither *fully* morphophonemic nor surface phonemic representations; instead, they are abstract codes influenced by orthographic features. For example in a syllable monitoring task, the participants found (perceived) a target syllable with a full vowel (/læg/; but not /lɔg/, in the control condition) in a word

---

[8] Foss and Blank (1980) maintain that to monitor a phoneme e.g., /t/, one must ignore the allophonic variations. On the other hand, they found that carrier word frequencies and the lexicality of carrier units did not have any influence on phoneme monitoring latencies. Therefore, they conclude that phonetic codes are abstract codes that do not include allophonic variations, and they are not the phonological (phonemic) codes stored in the lexicon.

containing a reduced vowel ('lagoon', /ləgu:n/) in the first syllable. Note that there is no morphologically related word to 'lagoon' which contains a full vowel in the first syllable. The only reason for such performance could be an orthographic influence on the phonological representation.

According to Taft & Hambly's (1985) and Chomsky & Halle's (1968) proposals, the realization of orthographic knowledge can be observed at phonological and at morphophonemic levels respectively. Thus, in tasks tapping into the acoustic or phonetic levels, no influence of orthography is expected.

1-4-2. Processing levels in phonological and metaphonological processing

Morais and associates (Morais, Kolinsky, & Claes, 1995) initially presented a two-level model for (meta)phonological processing of speech in which the identification of speech signals ends in conscious percepts of speech which become in turn the input to intentional phonological processing, i.e., metaphonological processing (Figure 5).

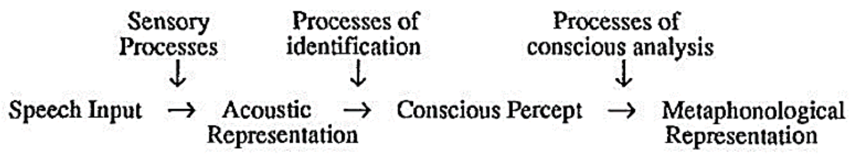


Figure 5. Two-level model presented by Morais and Collogues for phonological processing of speech (Morais, Kolinsky, & Claes, 1995), p. 201.

In this model, there are two levels dealing with the perceptual (identification) and conscious (metaphonological analysis) processing of speech. Based on this model, the tasks of identification and recognition tap into the perceptual processing level, whereas tasks such as segmenting into phonemes tap into the metaphonological processing level of speech processing.

Later, these authors found that not all tasks involving phonological processing can be accommodated in this model. For example, phoneme monitoring tasks[9] which

[9] In a phoneme monitoring task, participants are to decide whether a spoken input contains a target phoneme.

are designed to reveal the perceptual processes involve matching operations. Also, a task such as the ABX discrimination task[10] requires perceptual as well as matching operations. In fact, identifying a stimulus word is different from identifying whether a stimulus word is identical to another or not. The former requires the activation of the corresponding entry in the lexicon, and the latter involves a matching operation in addition to the perceptual processing.

Thus, perception and identification operations should be related to two different levels in speech processing. This newly introduced level is the identification/recognition level which can be influenced by knowledge from other sources such as orthography and by attentional processes that are not necessarily intentional (Morais, Castro, Scliar-Cabral, Kolinsky, & Content, 1987). This level is localized between the perceptual processes and the intentional analysis of speech (Figure 6).

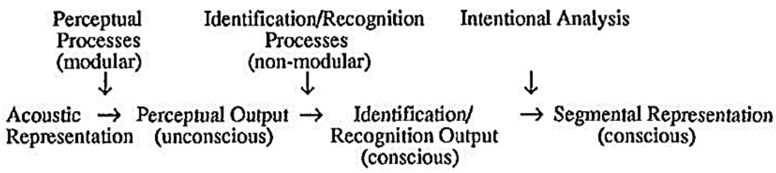


Figure 6. Three-level model presented by Morais and Collogues for phonological processing of speech (Morais, Kolinsky, & Claes, 1995), p. 208.

Thus, in this proposal, there are three levels (Figure 6 from left to right): perception, identification-recognition, and the intentional or formal explicit analysis (see (Kolinsky, 1998) for more empirical evidence for the three-level model). According to the authors, the perceptual level is immune to the orthographic knowledge, whereas the other levels may be influenced by orthographic knowledge.

1-4-3. Processing levels in the Representational Redescription Model

Karmiloff-Smith (Karmiloff-Smith, 1995) introduced a proposal about levels of cognitive processing under the name of Representational Redescription (RR) Model. Redescription is defined as recoding information stored in one representational format

[10] In the ABX discrimination task (auditory), the participants are to decide whether the third stimulus in a triad is identical to the first or second stimulus in that triad.

into a different one. She breaks the dichotomy of implicit/explicit representations and assumes that knowledge can be represented in four different formats: implicit (I), explicit 1 (E1), explicit 2 (E2), and explicit 3 (E3). The emergence of these formats is dependent on 'behavioral mastery' or receiving positive feedbacks, rather than on the developmental age. This model was not developed exclusively for phonological processing of speech; but it can be well applied to that.

At the 'I' level, representations are in a procedural format that is used to respond to stimuli in the external environment. This representational format is bracketed and not available to inter- and intra-domain operations in the cognitive system. The outcome of a procedure, but not its components, is available to other operators as a whole; thus, the representations are implicit at this level. Implicit representations make it possible to respond inflexibly (not suitable for other purposes), rapidly, automatically, and effectively to the environment.

Just by representational redescription, procedure components can become accessible for inter- and intra-domain operators at the E1 level. The redescription of representations does not mean that the 'I' representational formats are no longer available, but all remain there for appropriate automatic and rapid operations. At this level (E1), the 'I' representational formats are redescribed into a compressed and abstract format in a higher-level language. The components of the procedures are available for further operations. Such redescription enables the cognitive system to make analogies between or among different pieces of knowledge. The E1 representational formats are more flexible, simpler, and less special as compared to the 'I' representational formats. In other words, the conceptual redescriptions, in contrast to perceptual redescriptions, are more productive. E1 representational formats incorporate explicit manipulable knowledge based on which 'theories'[11] (or concepts) can be built and developed. The relationship between components is encoded at the E1 level; however, the E1 representational formats are not available to consciousness and verbal

[11] When the 'I' format is redescribed into the E1 format, the cognitive system begins to be flexible, and the person can build theories based on the E1 format (p.21). Theory here refers to how to exploit the stored knowledge to solve problems (Karmiloff-Smith, 1995).

reports. Although Karmiloff-Smith does not exclusively refer to speech processing and the role of orthography in speech processing, it can be understood that the E1 representational formats in the phonological domain may be influenced by knowledge from other domains (e.g., orthographic knowledge), since the E1 representational formats are not bracketed, and they are available for inter- and intra-domain representational links.

At the E2 level, representations are accessible to consciousness but not to verbal reports. The codes at the E2 and E1 levels are similar (spatial codes at the E1 level are redescribed into spatial codes at the E2 level which are available to consciousness).

At the E3 level, the knowledge becomes available for verbal reports and can be communicated. Knowledge from different domains can be linked when they are redescribed in a similar format at the E3 level.

The RR model assumes that knowledge can be represented at multiple levels that differ in the level of abstraction. Each level is associated with a certain format of representation. That is, for the same linguistic entity, such as a phonological unit, there can be multiple representational formats each of which appears at a certain level. Different methods used to elicit the knowledge (e.g., knowledge about a phonological unit) can tap into different representational formats. For example, perceptual tasks may tap into the implicit or procedural format, whereas metaphonological tasks may tap into the explicit formats. The interesting aspect of this model is that it breaks the dichotomy of implicit-explicit. Even among metalinguistic tasks, those that require verbal reflections tap into the most abstract representational format; namely the E3 representational format.

Explicit representations are developed gradually and enable one to reflect and analyze knowledge (Karmiloff-Smith, 1995). In her proposal, Karmiloff-Smith does not deal exclusively and explicitly with levels of speech processing, but her model can be well applied to levels of speech processing. For example, Morais and Kolinsky (Morais & Kolinsky, 2002) found that illiterates can tell that ‘pat’ is different from ‘pet’; but they cannot identify the segment which is different in them. Or, the same authors found that illiterates cannot delete the initial segments of spoken CVC syllables, whereas ex-illiterates can perform this task comparatively well. Such evidence may indicate that illiterates have and use procedural or E1 representational formats to distinguish the minimal pairs from each other; however, they might not have developed the corresponding E2/E3 formats.

Hence, literacy can be considered as a promoter for the representational redescription into explicit formats. As mentioned above, Karmiloff-Smith considers “behavioral mastery” as the factor that triggers representational redescription, but she does not discuss factors that might specifically lead to redescription in the language domain. Evidently, behavioral mastery in the domain of language perception/production does not lead to the development of the explicit phonological formats in illiterates (Morais, 2021; Morais, Bertelson, Cary, & Alegria, 1986). Goswami (2000) fills this gap and refers to the factors that lead to representational redescription or restructuring of phonological representations. These factors are discussed in the next section.

To summarize, some of the proposals discussed so far entail that orthographic knowledge could influence speech processing at some processing level(s) but not others. Orthographic knowledge could influence the phonological processing level, but not the acoustic and phonetic levels, in the proposal by Taft and Hambly. The proposal by Morais and associates limited the potential orthographic effect to the levels of identification/ recognition and intentional analysis of speech. The proposal by Karmiloff-Smith presents a detailed picture of representational formats, but it does not explicitly tell anything about possible effects of orthography on these representational formats.

In general, two of these proposals dealt with ‘where’ the orthographic effect has been/may be observed in speech processing. ‘How’ (in what way) orthographic knowledge can influence speech processing is something that is not explained in these proposals. The following section deals with how orthographic knowledge can influence speech processing and what factors are important in this process.

1-5. The nature of the orthographic effect on speech processing

Goswami (2000) proposes factors that are relevant in restructuring of phonological representations. She maintains that lexical phonological representations are restructured several times in the course of development, and that, over time, they become segmental and specific in terms of phonetic features (fully specified). According to this hypothesis, phonological representations of words are holistic in the early ages (see (Walley & Flege, 2000) for the same argumentation). By learning more words, the lexical phonological representations begin to be restructured from a global form to smaller units such as syllables and finally to phonemes (Fowler, 1991). Such phonological restructuring depends on the vocabulary size, word frequency, linguistic factors (e.g., 'sonority profile'), and word neighborhood density[12] (Figure 7). Words with many phonological neighbors and high frequency undergo phonological restructuring earlier than words with sparse neighborhood density and low frequency. Thus, the emergence of smaller phonological units is supposed to be the result of vocabulary growth and increasing experience with spoken language.

However, reading acquisition studies show that the emergence of phonemic representations is to some extent dependent on learning to read/write in an alphabetic writing system (e.g., (Morais, 2021; Goswami & Bryant, 1990)[13]). These authors assume that preliterates develop phonological representations at the level of syllables and rhymes. The restructuring in preliterates develops implicit or epilinguistic phonological representations that are used to recognize whether spoken words share some phonological units, but they are not sufficient to identify and produce those phonological units upon request (Gombert, 1992). Hence, according to this hypothesis, the emergence of phonemic representations (segments) depends partly on literacy.

[12] The number of words that are different from a target word in one phoneme (Ziegler, Muneaux, & Grainger, 2003).

[13] Morais (Morais, 2021) maintains that "phoneme is the conceptual heritage of the alphabetic literacy".

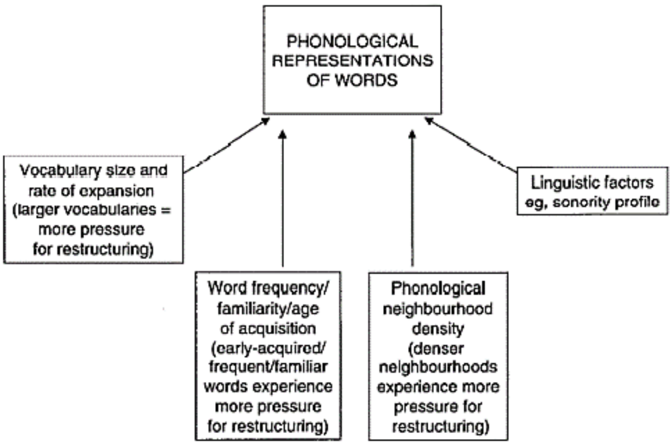


Figure 7. Important factors that influence phonological restructuring of words before literacy (Goswami, 2000), p.135.

Breaking the integrity of words into phonemic segments is something that happens as the result of literacy in an alphabetic script or similar trainings. In other words, graphemic representations in alphabetic scripts aid in representing phonemes (Figure 8).

According to this perspective, the transparency of the orthographic system that children learn to read and write in, is expected to have a role in segmental restructuring and representing phonemes. In transparent orthographies such as German, segmental restructuring occurs faster, and phonemic representations are acquired faster than in opaque orthographies. In other words, the degree of spelling-to-sound consistency (feedforward consistency) and sound-to-spelling consistency (feedback consistency) influences the degree of phonemic restructuring that occurs developmentally (Figure 8).

Fowler (1991) and Walley (1993) believe that lexical phonological restructuring and the development of phonological awareness are tightly interrelated. The implicit phonological representations for spoken units should be there for awareness to those phonological units to take place. Phonological awareness develops in stages; i.e., awareness to syllables and rhymes occur earlier than awareness to phonemes cross linguistically (Wimmer, Landerl, & Schneider, 1994; Cossu, Shankweiler, Liberman, Katz, & Tola, 1988; Liberman, Shankweiler, Fischer, & Carter, 1974). Also, awareness to syllables and rhymes are assumed to be universal, whereas awareness to phonemes is dependent on alphabetic literacy (Goswami, 2000).

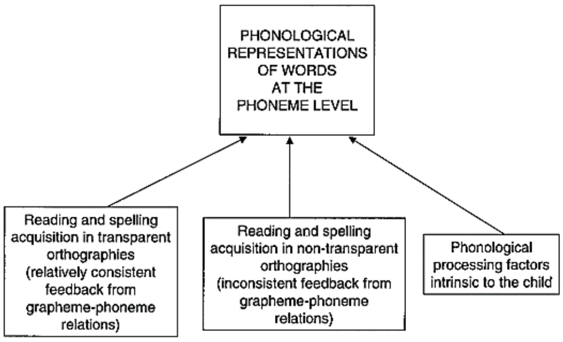


Figure 8. Important factors that influence phonological restructuring of words at the level of phoneme after literacy (Goswami, 2000), p.138.

The role of literacy/orthography in the emergence and development of phonemic representations is not limited to the explicit representations of phonemes used in metalinguistic tasks (phonemic awareness or phonemic representations used in reading/writing). There are findings suggesting that illiterates are less accurate than literates in recognizing whether two spoken words or nonwords sound the same or not (Nayernia, van de Vijver, & Indefrey, 2019), suggesting that literacy may enhance the precision of phonemic boundaries (Kolinsky, et al., 2021). Alternatively, the perception of phonemes represented consistently by the same letter/grapheme (e.g., /p/ represented consistently by <p> in English) may differ from the perception of phonemes represented by different letters/graphemes (such as /s/ that can be represented by <s> and <c> in English) (e.g., (Ziegler & Ferrand, 1998; Treiman, 1993; Goswami & Bryant, 1990); also, see Chapter 2). Thus, it can be assumed that not only explicit representations of phonemes but also the development and nature of implicit phonemic representations might be dependent on literacy and on the properties of a language's orthographic system.

In sum, literacy/orthographic knowledge can influence speech processing by restructuring of the phonological representations in terms of the orthographic features of languages. In this case, phonological representations for consistent words are supposed to be fully specified and finer grained as compared to phonological representations for inconsistent words (Ziegler, Muneaux, & Grainger, 2003) or as compared to words including phonological units without orthographic realizations (Van Ooijen, Cutler, & Norris, 1992). In other words, the nature of phonological representations may change in terms of orthographic features of a language. Thus, orthographic knowledge can restructure, specify, and reorganize phonological knowledge according to this account.

There is also another assumption about how orthography can influence speech processing. It might be the case that literacy/orthography exerts its effect on speech processing through simultaneous activation of cross-modal representations. Mental representations for events in one modality can co-activate the corresponding mental representations for the same event in other modalities after developing a cell assembly ((Hebb, 1949), for the application of the notion of Hebbian assembly to the orthographic consistency effect, see (Ziegler, Ferrand, & Montant, 2004) below).

Even mental imagery or a stimulus in one modality can co-activate the corresponding mental representations in another modality (Spence & Deroy, 2013). For instance, vanilla smell or hearing the word 'vanilla' may evoke the taste of sweetness or sweet foods in Western population (Spence, 2008)[14]. This phenomenon might emerge as a result of frequent co-occurrence of events in cross-modal activities, such as the co-occurrence of the vanilla smell (or the word 'vanilla') and tasting sweet foods. It has been shown that graphemes are one of the most common inducers in this respect (Mroczko-Wasowicz & Nikolic, 2014).

The cross-modal co-activation of mental representations can be either automatic or strategic i.e., conscious (Spence & Deroy, 2013). A mental representation that is "regularly" and "unavoidably" related to a mental representation in another modality is co-activated automatically ((Spence & Deroy, 2013), p.166). Alternatively, it might be the case that the link between the cross-modal events or representations trains the consciousness in a way that awareness is switched from one modality to another (Spence & Bayne, 2015) possibly to help the identification of a stimulus.

Given that literacy/orthography has been shown to influence speech processing, it can be assumed that links are formed between letters and sounds or between larger

[14] Synesthesia refers to a phenomenon "when additional perceptual experiences are elicited by sensory stimuli or cognitive concepts" (Hochel, et al., 2008).

units such as syllables or rhymes across modalities during reading and spelling. The corresponding cell assemblies lead to the cross-modal co-activation of mental representations (Ziegler, Ferrand, & Montant, 2004; Van Orden, 1987). Such a cell assembly is formed after a long exposure to scripts. For example, coming across the visual representation <l>, the corresponding phonological representation /l/ would be co-activated. Also, coming across the phonological representation /l/, the corresponding visual representation <l> can be co-activated.

Any inconsistency or mismatch in such links may create a competition or gap which breaks the normal process involving consistent links (one-to-one links) between a phonological representation and its corresponding visual representation (Ziegler & Ferrand, 1998). Thus, coming across a sound/phoneme such as /s/, multiple orthographic representations are assumed to be co-activated simultaneously: <s> and <c> in English. This inconsistency in the link between the phonology (auditory) and orthography (visual) can hinder the normal processing taking place for a phoneme such as /l/ in English.

This cross-modal activation of mental phonological and orthographic representations can also be either automatic (Ziegler & Ferrand, 1998) or strategic (Cutler, Treiman, & Van Ooijen, 2010). If it is automatic, the processing problem should occur whenever an inconsistency is met and at all levels of processing i.e., lexical and sub-lexical. If the co-activation of cross-modal representations is strategic, attentional factors as well as the existence of the inconsistency are expected to lead to a problem in processing. For instance, including a large number of inconsistent trials in the stimuli list can draw the attention toward the inconsistency in the link between phonology and orthography which results in the strategic application of orthographic knowledge in addition to phonological knowledge in an auditory task (Cutler, Treiman, & Van Ooijen, 2010). A strategic application of orthographic knowledge should occur at higher levels of speech processing (post lexical level) where the orthographic information is available. By contrast, if the same inconsistency is presented without drawing attention to it, it should be processed like a consistent case.

To summarize, a role for orthography in speech processing has not been incorporated in all models of spoken word recognition. Just the BIAM (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005) assumes an explicit and automatic role for orthographic knowledge during speech processing at lexical and prelexical levels. The Merge model predicts a strategic effect of orthographic knowledge in speech processing at higher levels of processing (postlexical) when orthographic knowledge is induced in auditory tasks, when it must be used to accomplish tasks (e.g., reading/writing), when it is rendered salient, or when it helps task performance (e.g., in highly cognitively demanding auditory tasks).

Thus, the existing accounts of spoken word processing differ considerably with respect to the question of whether and how it can be affected by orthography. At the same time, as it has been mentioned, the cross-linguistic database is small in this respect (Goswami, 2000; Walley & Flege, 2000), and more investigations are required to have a clearer picture of whether and in what precise way orthographic knowledge can influence spoken word processing.

This study is an attempt to extend the cross-linguistic literature in this respect. It is of interest to know whether and how orthographic knowledge can influence spoken word processing in Persian. Comparing the peculiarities in Persian and other languages studied so far for that purpose, it becomes clear that the mapping pattern between phonology and orthography is different in Persian as compared to other languages studied so far (mainly French, English, and Portuguese). In the following section, these properties of Persian are briefly explained.

1-6. An overview on Persian phonology and orthography

Persian, alternatively called Farsi, is an Indo-Iranian language belonging to a sub-branch of the Indo-European family (Lewis, 2009). This language is the official language in Iran and is spoken in some part of Afghanistan (called Dari) and Tajikistan (called Tajik) as well (Toosarvandani, 2004). The Persian language can developmentally and historically be divided into three stages: Old Persian ca. 550 - 330 B.C., middle Persian 330 BC -652 A.D., and modern Persian 625 AD - present (Natel Khanlari, 1987). The Persian language has been written in three different scripts through its history. Respectively from the oldest to the current, these are Cuneiform, Pahlavi (since Sassanid Empire), and modified Arabic script, ca. since 652 A.D.

(Daniels & Bright, 1996; Natel Khanlari, 1987). The focus of this study is on the Modern standard Persian language spoken in Iran which is written in the modified Arabic script.

Modern Persian has 6 vowels: /i, u, e, o, æ, ɑ/ (Samareh, 1985), which are divided into two categories based on quantitative and qualitative features. Based on the quantitative feature, one mora is assigned to the short vowels /æ, e, o/, and the long vowels are dimoraic /ɑ, u, i/ (Hayes, 1989). This classification is important in prosodic structures associated with weight units (Gordon, 2004); namely, the long and short vowels are given respectively two and one weight unit(s) in the prosodic structure, i.e., in versification (Deo & Kiparsky, 2011; Lazard, 1992). Qualitatively, the Persian vowels fall into three categories based on the height of the vowels (the most prevalent categorization): high /i, u/, mid /e, o/, and low /æ, ɑ/ from which the first member of each pair is a front vowel /i, e, æ/, and the second member of each pair is a back vowel /u, o, ɑ/ (Samareh, 1985).

From these two features (quantitative and qualitative), the second one is argued to be the prevalent one for the vowel inventory of the Modern Persian (Rohani-Rahbar, 2012; Tsukada, 2011). For example, Kramsky (1966) reported that the duration of the vowels in e.g., /ɣ**o**m/ and /n**ɑ**m/ are the same 277 *ms*; although the former is categorized as a short vowel and the latter as a long vowel (see also (Samareh, 1985) for the same argument). Moreover, there is empirical and theoretical evidence (e.g., stress pattern) that shows that length is not contrastive in Modern Persian (see (Rohani-Rahbar, 2012; Tsukada, 2011)).

It is believed that a restructuring has occurred in the Persian vowel system after the Old and Middle Persian (written in Cuneiform and Pahlavi scripts). In these two eras, the vowels contrasted in quantity (with eight vowels), whereas in Modern Persian, the vowels do not contrast in quantity anymore, but contrast in quality instead (Natel Khanlari, 1987; Windfuhr, 1979). It is yet to be investigated whether the deployment of the modified Arabic script for writing Persian had any impact on the simplification of the Persian vowel system. Although the system underwent a restructuring, the labels of 'short' and 'long' are still used for vowel categorization in Modern Persian in which length is no more contrastive (Samareh, 1977). Therefore, also in the present investigation, I use these labels.

The short vowels /æ, e, o/ are not orthographically represented in the conventional writing of Modern Persian (Daniels & Bright, 1996). For instance, a word including a short vowel, such as /ɣæm/ is written as <ghm>. In the early stages of learning to read and write (first grade of elementary school), short vowels are orthographically represented by diacritics that are later removed from the script. /æ/ is represented by <◌َ>; /o/ is represented by <◌ُ>; /e/ is represented by <◌ِ> in the 'pointed' version or the unconventional writing system. In contrast, long vowels are orthographically represented in the conventional writing of Persian by letters: /ɑ/ <آ>, /u/ <و>, /i/ <ی>. Therefore, short vowels contrast with long vowels in that short vowels have no orthographic representation, whereas long vowels do. It can be said that the mapping between phonology and orthography is incomplete in the case of Persian short vowels (i.e., mismatch between phonology and orthography), since short vowels are not orthographically represented. By contrast, the mapping is complete for Persian long vowels. This contrast is the focus of this dissertation.

There are twenty-three consonants in Modern Persian - henceforth Persian - (Windfuhr, 1990). Some consonants are represented by multiple graphemes (inconsistent sound-to-spelling mapping), and others are always represented by the same letter (consistent sound-to-spelling mapping). For example, /m/ is consistently written as <م>; however, /z/ is written with either of these graphemes <ز, ض, ذ, ظ>. It is completely arbitrary but conventional which grapheme should be used to represent an inconsistent phoneme in a word (Baluch, 2005). For example, the phoneme /z/ is written as ز in زر <zr>, /zar/ (gold); but, as ض in ضرر <zrr>, /zærær/ (damage).

In total, five consonants are polygraphic (represented by different graphemes), /t, s, h, z, ɣ/ <ط - ت, ث - ص - س, ح - ه, ذ - ز - ظ - ض, غ - ق>, and the rest are consistently represented by the same graphemes. The contrast between inconsistent and consistent consonantal phonemes is also in the focus of this study. Note that for the five consonants with inconsistent sound-to-spelling, the spelling-to-sound relation is nonetheless consistent (e.g., /ɣ/ can be spelled by <ق> or <غ>; but <ق> and <غ> represent always the same sound /ɣ/).

The realization of some of the graphemes depends on the connections they have on the sides. For example, /z/ <ض> is realized as <ضـ> if another grapheme is attached to the left of it (e.g., in ضرر), or if it is connected on both sides (e.g., in مضر). But, if another grapheme is attached only to the right of it, or if it has no attachment on any side, it is realized as <ض> such as in مریض or مرض (Persian is written from right to left). This is some sort of capitalization where <ض> is the capital letter, but <ضـ> is the small letter for the grapheme <ض> /z/. According to Dehaene and colleagues (Dehaene, Cohen, Sigman, & Vinckier, 2005), the font, size, and capitalization have no influence on written word recognition; therefore, it is assumed that ض is recognized independently of its surface realization as ض or ضـ.

Persian has three syllable structures: CV (as in /tɑ/ 'fold'), CVC (as in /ɣæm/ 'sorrow'), or CVCC (as in /mærg/ 'death'); thus, the form of syllables is CV(C)(C) in Persian. Stress falls on the final syllable of a word regardless of the vowel category (short or long) in that syllable (Kahnemuyipour, 2003).

1-7. Testing for influences of orthography on speech processing in Persian

As explained above, there are two features of the relationship between Persian orthography and phonology that lend themselves well to test the influence of orthography on speech processing. These two features are in the focus of this dissertation: first, the incomplete match between phonology and orthography in the case of short vowels (vs. long vowels); second, the inconsistent match between phonology and orthography in the case of the polygraphic (vs. monographic) consonants.

1-7-1. Incomplete match between phonology and orthography

As mentioned earlier, short vowels have no orthographic representations in Persian conventional writing, whereas long vowels do. This feature was used to investigate whether Persian stimulus items containing short vowels were processed differently from stimulus items containing long vowels when presented auditorily. If they were, this could be considered as evidence for an influence of orthography on speech processing.

To this aim, four experiments were administered, each tapping into a certain level of processing. Across experiments, it was assumed that if orthography had an influence on speech processing, processing (pseudo)words containing short vowels should be less accurate than processing (pseudo)words containing long vowels when the task was performed by Persian literates. An auditory AX discrimination task and a phoneme reversal task were administered to compare the performance for (pseudo)words containing short vowels with the performance for (pseudo)words containing long vowels. It was assumed that these two tasks tapped into implicit and explicit processing of phonological information respectively. It was hypothesized that if Persian orthography had an influence on speech processing, pairs of Persian (pseudo)words containing short vowels such as /sæm-som/ written as <sm>, <سم> should be processed (discriminated) with higher error rates than pairs of Persian (pseudo)words containing long vowels such as /kɑh-kuh/ written as <kah-kuh>, <کاه-کوه> by Persian literates. However, if orthography had no influence, (pseudo)words containing short vowels were expected to be processed as accurately as (pseudo)words containing long vowels. In the same vein, Persian literates were expected to manipulate (reverse) short vowels with a higher error rate compared to the long vowels in the auditory phoneme reversal task. The performance in this task could be an index for the phonological awareness to short and long vowels.

Two control groups were asked to participate in these experiments to make sure that any possible differences between the performances on Persian (pseudo)words containing short or long vowels were exclusively due to the difference in their orthographic representation. Persian illiterates were expected to show the same performance for (pseudo)words containing short or long vowels, because they would not have orthographic representations. German literates were also expected to show the same performance for (pseudo)words containing short or long vowels as both short and long vowels are orthographically represented in the German writing system. Therefore, a German AX discrimination task and a German phoneme reversal task were added. Persian literates and German literates participated in both versions of the AX discrimination task. Due to the tasks used in the first two experiments, the results we obtained could not be interpreted unambiguously, because it was not entirely clear how the tasks were performed. More specifically, it was possible that the AX discrimination task could have been performed based on strictly perceptual representations, and such representations would be resistant to any influence of orthography (Morais, Kolinsky, &

Claes, 1995). For the phoneme reversal task, we could not rule out the application of orthographic representations instead of or in addition to phonological representations. Hence, we decided to confirm our results by using two other experimental paradigms requiring implicit and explicit processing of phonological information while minimizing the chance of relying exclusively on perceptual or orthographic representations to the extent possible. An auditory ABX discrimination task was supposed to tap into a higher level of phonological processing because it included a matching component besides a discrimination component. It was assumed that matching auditory inputs would force the participants to operate on a higher processing level and go for identification/ recognition processing of the stimuli rather than purely perceptual processing (Morais, Kolinsky, & Claes, 1995).

In the Persian auditory ABX discrimination task (implicit phonological processing), Persian literates and illiterates were instructed to decide whether a third stimulus was identical to the first or to the second stimuli in each trial. For example, they were to decide whether /ʤerm/ was identical to the first or the second item in the /ʤerm-ʤorm/ pair.

In the Persian phoneme monitoring task (explicit phonological processing), Persian literates and illiterates were to decide whether auditory (pseudo)words contained a target phoneme or not. In other words, they were to match the segments of the auditory inputs against the target phoneme to find whether the auditory input contained the target phoneme. In both tasks, the performance of Persian literates on (pseudo)words containing short vowels was expected to be more error prone than the performance on (pseudo)words containing long vowels. Persian illiterates were not expected to show such a performance difference.

Additionally, two reading tasks, naming written pseudowords with long vowels or pointed short vowels and semantic categorization of real words written in the conventional or the pointed script, were conducted to find whether performance in these tasks correlated with performance in the auditory ABX task. According to one account of the effect of orthography on speech processing ('*offline*' account), prolonged exposure to Persian conventional writing leads to a restructuring of phonological representations in terms of orthographic features. In proficient readers of Persian conventional writing such restructuring based on the lack of orthographic representation for short vowels might be more pronounced and in consequence they might show a stronger orthographic effect in the auditory ABX discrimination task. Conversely, proficient readers of Persian unconventional writing (pointed version) might be expected to show less restructuring and hence a weaker orthographic effect in the auditory ABX discrimination task.

1-7-2. Inconsistency between phonology and orthography

Some consonants are polygraphic in Persian; that is, they are spelled in different ways. Such cases are referred to as 'inconsistent cases'. In contrast, other consonants are spelled in one way only. Such consonants are referred to as 'consistent cases'. Therefore, spelling consistency is another point of contrast in this investigation. If orthographic knowledge had an influence on speech processing, the spoken words with inconsistent spelling were expected to be processed slower than the spoken words with consistent spelling.

An auditory primed lexical decision task was conducted to examine the orthographic consistency effect in Persian. It has been shown that phonological priming has a facilitatory effect on the recognition of spoken targets (Dufour & Peerman, 2004; Slowiaczek, Nusbaum, & Pisoni, 1985). When targets are primed by phonologically related words, they are recognized faster than when they are primed by phonologically unrelated words. This facilitation can be modulated by orthographic consistency. Jakimik and colleagues (Jakimik, Cole, & Rudnicky, 1985) observed a facilitatory effect of phonological priming when corresponding spoken primes and targets were consistently spelled (message – mess) but not when corresponding spoken primes and targets were spelled inconsistently (definite – deaf).

The consistency effect has been reported for English, a language with inconsistencies in both sound-to-spelling ('deal, peel') and spelling-to-sound ('deal, deaf') correspondences (Slowiaczek, Soltano, Wieting, & Bishop, 2003). It was of interest to know whether the orthographic consistency effect could also be found in a language such as Persian with inconsistent sound-to-spelling correspondences but almost consistent spelling-to-sound correspondences in the case of consonants.

If orthographic consistency would modulate the facilitatory effect of phonological priming, that effect might be absent or attenuated by orthographic inconsistency. Hence, lexical decision latencies for inconsistent targets were expected to be the same when primed by phonologically related words (as in /t̲ænɑb-**t̲æn**/ <طناب - **تن**> combination of prime-target) as when primed by unrelated word forms (as in /kuʧeh-**tæn**/ <کوچه - **تن** >). By contrast, lexical decision latencies for consistent targets were expected to be faster when primed by phonologically related words (as in /k̲uʧeh-**k̲uʧ**/, <کوچه - **کوچ**>) than when the same targets were primed by unrelated word forms (as in /pedær-**kuʧ**/, <پدر - **کوچ**>).

1-8. Organization of the thesis

In this first chapter, the question of possible influences of orthography on speech processing has been introduced, and theoretical foundations of the empirical study have been presented. As the study is on Persian, relevant features of the Persian language have been explained briefly to show how speech processing might be affected by the mapping pattern between phonology and orthography in this language.

In the next chapter (Chapter 2), an overview of the related literature will be presented. This chapter is divided into three sections. In the first section, some studies are presented that investigated a possible influence of orthography on explicit phonological processing of speech (metaphonological processing). In the second section, some studies are presented that investigated a possible influence of orthography on implicit phonological processing of speech. The last section deals with the influence of literacy on phonological processing.

In the subsequent chapters, the experiments I conducted, and their results are presented and discussed in the order of implementation:

- The third chapter reports the Persian/German AX discrimination experiments and the Persian/German phoneme reversal experiments performed by Persian literates, Persian illiterates (participated in the Persian AX discrimination task only), and German literates.
- In the fourth and fifth chapters, the Persian auditory ABX discrimination experiment and the Persian phoneme monitoring experiment are reported.

- In the sixth chapter, the reading experiments are reported. In addition, I will report a correlational analysis between the performance in the ABX discrimination experiment and the performance in the reading experiments.
- In the seventh chapter, the auditory primed lexical decision experiment is reported.
- A summary of the results and the interpretation of the data are presented in Chapter 8, Summary and general discussion.
-Chapter 9 includes the key findings of the dissertation and the theoretical implication of them for models of word processing. This chapter ends in pointing out the limitations of the study and some notes on further investigations.

Chapter 2: A review of the literature on the influence of orthography on speech processing

2-1. Introduction

The term 'literacy' may invoke several meanings and interpretations. The intended interpretation in this dissertation is the one used by cognitive psychologists: the ability to read and write (Morais & Kolinsky, 2002). By literacy, a new system of representations, i.e., orthographic representations, is introduced into the mental system, and processing orthographic representations may interact with other already-existing mental representations or processes. The comparison of illiterate[15] and literate populations provides an opportunity for researchers to find relationships between literacy and cognitive developments in normal adults. To name just one example: It has been shown that literates are more accurate than illiterates in recognizing two dimensional pictures (Petersson, Reis, & Ingvar, 2001). Reading ability correlates strongly with other cognitive abilities[16] even when covariates such as age, sex, education, and health status are controlled (Barmes, Tager, Satariano, & Yaffe, 2004). Comparing the performance of literates and illiterates can help identifying influences of orthography on spoken language processing.

The system developed to process orthographic representations is not an isolated and independent system; but it makes connections to other already-developed systems engaged in processing other types of representations such as phonological representations. Through these connections, the structure of the already-established knowledge and the functionality of other systems may incur some changes; thus, it is important to know how literacy affects cognition including speech processing.

Moreover, the system of representations or codes that are introduced to mind and cognition through literacy varies cross linguistically. In some languages, the codes represent syllables (e.g., in Japanese); in other languages, the codes represent

---

[15] Illiterates are those normal population (without any cognitive impairments) who could not acquire reading and writing due to some sociocultural factors. These factors include 'literacy is not necessary, illiterates should take care of others in the family, no school was available in the region/village, illiterate should work instead of attending the literacy course, or parents had low income'.

[16] Cognitive ability was evaluated by cognitive functions such orientation to time and place, verbal memory (words recall ability), attention (ability to focus on a task), calculation, language ability (naming different words starting with a given phoneme), and visuospatial ability.

morphemes (e.g., in Chinese); yet in others, the codes represent the sounds/phonemes of the language (e.g., in English). The focus and concern of this dissertation is on the third type which is called alphabetic literacy. Compared to the other types (syllabic literacy and logographic literacy), the codes in the alphabetic writing systems present the smallest segment of speech, i.e., sounds/phonemes (Gelb, 1952). Thus, alphabetic literacy may require a higher analytic faculty; hence, it is expected to have a higher impact on cognition (Morais & Kolinsky, 2002).

Even in languages with alphabetic system, the mapping between phonology and orthography varies in opacity, which may lead to a graded influence of orthography on phonological processing across alphabetic writing systems. Comparing the performance of literates in transparent orthographies against the performance of literates in opaque orthographies is another way to examine the influence of orthography on phonological processing. Alternatively, comparing the performance of literate speakers of a language where the links between phonology and orthography are consistent, incomplete, or inconsistent can help find how orthography can influence speech processing in the speakers of a language.

The present investigation is devoted to study the influence of literacy or orthographic information on spoken language processing, specifically on the phonological processing of spoken language. Researchers in this field have made a distinction between two forms of phonological processing: phonological processing which is part of "knowing language" and phonological processing which is part of "believing about language" that are respectively related to the language and metalanguage competencies (Morais & Kolinsky, 2002). To illustrate this distinction: illiterates can discriminate minimal pairs; they are, however, unable to identify different segments in minimal pairs or manipulate phonemes (Scliar-Cabral, Morais, Nepomuceno, & Kolinsky, 1997). According to Scliar-Cabral et al. (1997), such observations indicate that two different types of phonological processing are involved: the phonological decoding to discriminate minimal pairs and the phonological processing needed to identify or manipulate phonemes. The former is supposed to be an implicit process (unconscious, part of knowing language), whereas the latter is considered to be an explicit process (conscious, part of believing about language).

The findings of a number of studies show that alphabetic experience correlates with the development of phonemic awareness[17], an explicit, metaphonological ability (e.g., (Lukatela, Carello, Shankweiler, & Liberman, 1995; Kolinsky, Morais, Content, & Carry, 1987; Morais, Cary, Alegria, & Bertelson, 1979)).

There are opposing views about the influence of orthography on phonological processing at the implicit level. Some findings suggest that phonological processing is not vulnerable to orthographic knowledge at the perceptual level, and the orthographic knowledge can influence phonological processing only at a post perceptual level (e.g., (Cutler, Treiman, & Van Ooijen, 2010; Morais & Kolinsky, 2002; Morais, Kolinsky, & Claes, 1995)). By contrast, other authors assume that orthographic knowledge can influence phonological processing at the perceptual level in *online* and *offline* manners (e.g., (Taft, 2011; Ziegler, Ferrand, & Montant, 2004)).

According to Fodor (1983), the perceptual processes through which the perceptual units are derived from acoustic inputs are modular and automatic. Therefore, the knowledge outputted from this stage can be reworked by orthographic knowledge at the post perceptual level only. In a dichotic listening paradigm to study the effect of literacy on phonological fusion, the participants (literates and illiterates) were asked to report what they heard while they received separate auditory strings in each ear simultaneously (unpublished experiment by Castro and Morais reported in (Morais & Kolinsky, 1995)). For example, the participants heard ‘back’ in one ear and ‘lack’ in another ear simultaneously. The inputs were expected to be illusorily perceived as ‘black’ with a consonantal cluster in the onset. The authors selected pairs of words whose phonological fusion outputs sounded like an existing word, but was consistent or inconsistent with its spelling. For instance, in the pair ‘pena-lena’, the phonological fusion output /plena/ was consistent with its spelling <plena>. But in the pair ‘par-lar’, the phonological fusion output /plar/ was inconsistent with the word spelling <pelar>. The rate of the correct phonological fusion output was high for the consistent condition (60%) and almost the same across the literates and illiterates. However, in the

[17] The ability to identify, manipulate and/or reflect on phonemes of a spoken language (Gombert, 1992).

inconsistent condition, the rate of the correct phonological fusion output was 55% and 17% across the illiterates and literates respectively. In the inconsistent condition, the drop in the rate of the correct phonological fusion output was significantly more pronounced in the literates than in the illiterates who did not consider any spelling restrictions. This difference could indicate that the perceptually perceived string 'plar' was reworked at the post perceptual level considering its orthographic knowledge by the literate group.

In contrast to researchers who limit the influence of literacy (orthography) to the post perceptual level, another group of researchers believe that the effect of literacy is not limited to the post perceptual level; instead, literacy can influence speech processing at the perceptual level through a top-down flow of information or through the activation of cross system codes. There are two accounts for the influence of orthography on speech processing at the perceptual level. According to the *online* account, hearing a spoken word activates phonological and orthographic codes for that spoken word simultaneously at the prelexical and lexical levels (Pattamadilok, Perre, Dufau, & Ziegler, 2009; Taft, Castles, Davis, Lazendic, & Ngye-Hoan, 2008; Ziegler, Ferrand, & Montant, 2004; Ziegler & Ferrand, 1998). When the links between the corresponding activated phonological and orthographic codes are consistent (i.e., one-to-one), the activation level of the phonological codes is reinforced by the activation of the corresponding consistent orthographic codes. In contrast, when multiple orthographic codes are activated for a given activated phonological unit in the inconsistent cases, the competition among the activated orthographic codes would slow down the spoken word processing. Thus, according to the *online* account, orthographic knowledge affects the perception of the speech in real time through direct links between phonology and orthography.

In the *offline* account of the influence of orthography on speech processing, it is assumed that literacy structures the phonological representations in terms of the orthographic features (Taft, 2011; Ziegler & Goswami, 2005; Goswami, 2000). Empirical evidence suggests that the border between phonemes may become sharper when the phonological representations are consistently spelled (Kolinsky, 2015; Van Ooijen, Cutler, & Norris, 1992). Also, phonological processing may be different at the prelexical level in literates and illiterates (see (Reis & Castro-Caldas, 1997) for the effect of literacy in a pseudoword repetition task). Thus, hearing spoken words would not simultaneously activate the corresponding orthographic codes, but through learning to read and write, orthographic representations have restructured the phonological representations. It has also been suggested that when phonological representations are consistently spelled, they get activated faster by raising the activation levels of phonological codes (Muneaux & Ziegler, 2004; Ziegler, Ferrand, & Montant, 2004).

Thus, the involvement of orthographic knowledge in spoken word processing has been reported for the perceptual (Ziegler & Ferrand, 1998) and post perceptual (Morais & Kolinsky, 1995) levels, as well as in tasks with metaphonological components (phoneme monitoring (Dijkstra, Roelofs, & Fieuws, 1995), phoneme counting (Treiman & Cassar, 1997), and phoneme deletion and reversal (Castle, Holmes, Neath, & Kinoshita, 2003)) and tasks without metaphonological component (lexical decision (Ventura, Morais, & Kolinsky, 2007), shadowing (Ziegler, Muneaux, & Grainger, 2003), semantic categorization (Peerman, Dufour, & Burt, 2009)). On the other hand, there are also some experiments whose findings did not show any influence for orthographic knowledge on speech processing (Damian & Bowers, 2009; Roelofs, 2006; Chen, Chen, & Dell, 2002). The details of such studies are discussed in the following sections.

The literature review on this topic is divided into three sections: the influence of orthographic knowledge on metaphonological processing, the influence of orthographic knowledge on phonological processing, and the effect of literacy on phonological processing.

2-2. Orthographic knowledge and metaphonological processing

Phonological awareness has been defined as the ability to perform intentional mental operations and manipulations on speech units (Morais, 1991; Tunmer & Rohl, 1991). Authors differ with respect to the units to which phonological awareness applies: phonological awareness as the awareness to syllables (Mann, 1991), phonological awareness as the awareness to phonemes (Tunmer & Rohl, 1991), or phonological awareness as syllable awareness, onset/rhyme awareness, and phonemic awareness (Treiman, 1991).

In the literature, phonological awareness has also been defined in terms of the task difficulty which is used to assess this ability. In this approach, factors such as the demands on the linguistic knowledge, on memory, and on the analytic competence have been incorporated (Morais, 1991). For example, tasks that require a classification of the phonological units (e.g., rhyme judgment: whether two spoken strings rhyme) are easier than tasks which require a manipulation of the phonological units (e.g., phoneme reversal: reversing the phonemes of spoken strings).

There is also no agreement in the literature about the emergence of phonological awareness. Some believe that it emerges at the age of 4-5 (Rosner & Simon, 1971), and others consider phonological awareness to emerge at the age of 6 (Bruce, 1964) or later (Morais, 1991). This difference is due to recruiting different methodologies or examining different levels of phonological awareness. For example, syllable awareness and rhyme awareness appear earlier than phonemic awareness (Morais, Bertelson, Cary, & Alegria, 1986). It is argued that since syllables are the salient unit in speech perception and production, syllable awareness emerges earlier than phonemic awareness. Phonemes, by contrast, are coarticulated in speech; hence, they are difficult to identify and manipulate (Liberman & Shankweiler, 1977). The sequential emergence of syllable and phoneme awareness has been observed in several languages (see e.g., (Morais, 2021; Cossu, Shankweiler, Liberman, Katz, & Tola, 1988)).

Treiman (1983) found an additional level between syllable awareness and phonemic awareness: onset/rhyme awareness. In the hierarchical structure of the syllable, onset/rhyme is a level between syllable and phoneme (Halle & Vergnaud, 1980). Awareness to rhyme has been shown to emerge before literacy: preliterate children are sensitive to rhyme (Bryant & Bradley, 1985). Rhyme judgement can be performed by considering the overall similarity of the stimuli. In comparative studies, the developmental order of syllable awareness, rhyme awareness, and finally phonemic awareness has been reported (Treiman & Zukowsky, 1991).

It is well accepted that literacy has an extensive influence on the metaphonological ability or phonological awareness (Kolinsky, 1998). About the relationship between phonological awareness and literacy, there are, however, different views. Learning alphabetic codes brings about some sort of metaphonological ability, because orthographic codes play an important role in creating the highly abstract phonological representations, i.e., phonemes (Gonzalez & Gonzalez, 1993). Thus, learning to read/write is supposed to be a prerequisite of phonological awareness. It has been, on the other hand, argued that children must first learn how to manipulate phonological units and only then their correspondences to graphemes (Alegria, 1985), hence phonological awareness is assumed to be the prerequisite of learning to read/write.

Correlational studies show a strong correlation between literacy and phonological awareness (Sebastian & Maldonado, 1986; Tunmer & Nesdale, 1985). Some studies consider the level of phonological awareness a predictor of the reading level: for example syllabic awareness as the best predictor of later reading ability (Mann & Ditunno, 1990); intra-syllabic (onset/rhyme) awareness as the best predictor of later reading ability (Bryant, Maclean, Bradley, & Crossland, 1990); phonemic awareness as the best predictor of later reading ability (Calfee, 1977); or intra-syllabic and phonemic awareness as the best predictor of later reading ability (Lundberg & Hoien, 1991).

In spite of a vast number of studies, it is still debated whether phonological awareness at the level of the phoneme is a prerequisite or the consequence of alphabetic literacy. The proponents of the prerequisite view suggest that phonemic awareness is the prerequisite of learning to read, and it facilitates reading (Lundberg & Hoien, 1991; Liberman, 1973). This is because learning to read requires manipulation of segments in speech; thus, children should be aware of these units at the time of learning to read. On the other hand, it has been shown that whereas pre-readers and illiterates are sensitive to syllabic (Maclean, Bryant, & Bradley, 1987) and intra- syllabic (Treiman & Zukowsky, 1991) structures of speech, their phonemic awareness was low (Morais, Bertelson, Cary, & Alegria, 1986). Even literates in non-alphabetic writing systems have been reported to be phonemically unaware (Mann, 1986). Based on these findings, it seems most plausible that phonological awareness at the level of the phoneme is the consequence of learning to read and write in alphabetic writing systems. In addition to these two perspectives (phonological awareness as the prerequisite or consequence of learning to read), there is a third view that assumes a reciprocal directionality between phonemic awareness and learning to read and write (Morais, 1991). Based on this view, reading instructions provide access to the highly abstract phonological representations (phonemes) which in turn enhances the ability to read and write. The empirical evidence is largely in favor of the 'consequence' and the 'reciprocal' perspectives (Morais, 2021; Kolinsky, 2015; Morais, Bertelson, Cary, & Alegria, 1986; Morais, Cary, Alegria, & Bertelson, 1979).

The results of a seminal study on illiterate and ex-illiterate adults showed that illiterates could not add a consonant to the beginning of spoken VC strings, also they could not delete the initial consonant from spoken CVC strings, whereas ex-illiterates had a reasonably high performance in these tasks (Morais, Cary, Alegria, & Bertelson, 1979). These results show that phonemic awareness does not appear spontaneously; rather, learning to read/write or some similar trainings has a major role in its emergence. Even though the emergence of syllabic awareness and intra-syllabic awareness is not dependent on literacy, these abilities nonetheless improve extensively with literacy (Bertelson, de Gelder, Tfouni, & Morais, 1989). Thus, phonemic awareness requires the explicit analysis or segmentation of spoken strings which is dependent on alphabetic literacy (Morais, 2021).

The influence of orthographic knowledge on metaphonological processing has not only been found in phoneme addition/deletion tasks with literates and illiterates. In phoneme counting (Ehri & Wilce, 1979) and phoneme reversal (Castle, Holmes, Neath, & Kinoshita, 2003) tasks, orthographic knowledge had an influence on the performance. In the auditory phoneme counting experiment by Ehri and Wilce (1979), literates detected one more sound in /pɪʧ/ than in /rɪʧ/ due to the presence of the silent letter in the former (<pitch>) but not in the latter (<rich>).

An influence of orthographic knowledge on auditory phoneme reversal has been reported by Castle and associates (Castle, Holmes, Neath, & Kinoshita, 2003). The phonemes of orthographically transparent words (such as 'mood') could be reversed more accurately than the phonemes of orthographically opaque words (such as 'gnome') by literate participants. Transparency, here, was defined as the mapping pattern between graphemes and phonemes. For example, in ‘mood’, there is a direct relationship between the sounds and corresponding graphemes. However, in ‘gnome’ this relationship is indirect: there are silent letters in ‘gnome’. Thus, as the results of this experiment showed orthographic knowledge and phonemic awareness seem to be closely interrelated.

It might be argued that metaphonological tasks, which require the manipulation of phonological units (e.g., phoneme addition/deletion, phoneme reversal, and phoneme counting), are difficult to perform. Therefore, participants may have used their orthographic knowledge to facilitate the task performance. However, the influence of orthography has also been reported in easy-to-perform metaphonological tasks that required only the detection of phonological units (i.e., where no manipulation of phonological units was required). For instance, rhyme judgment[18] was faster for the spoken pairs such as ‘toast-roast’ as compared to ‘toast-ghost’ in a rhyme judgment task performed by literate participants. In the former pair, the rhyming words were orthographically consistent; while the rhyming words in the second pair were orthographically inconsistent (Seidenberg & Tanenhaus, 1979). Orthographic knowledge thus seems to have influenced the performance in a task that was not difficult to perform and did not require the manipulation of speech segments.

Similarly, Halle and colleagues (Halle, Chereau, & Segui, 2000) found that French literate speakers had significantly more problem in detecting /p/ in ‘absurd’ /æpsyrd/ because /p/ is spelled as <b> in this word; whereas, /p/ was detected easily in ‘lapsus’ /lapsys/ in which /p/ is spelled as <p>. Thus, the detection of speech sounds was influenced by orthographic knowledge.

The influence of orthography has also been reported for analysis of spoken units larger than rhymes or phonemes. In an auditory blending task (Ventura, Kolinsky, Brito-Mendes, & Morais, 2001), Portuguese literate participants were asked to blend a

[18] In rhyme detection, the participants are to judge whether auditory items have the same body (-VC) sounds.

first spoken CVC string into a second spoken CVC string. The strings ended either in consonants, such as ‘bar-mel’, or in a silent letter <e>, such as ‘cure-pele’. For pairs such as ‘bar-mel’ ending in consonants, blending occurred in the form of C/VC (/bel/). However, for pairs with silent ultimate ‘<e>’ such as ‘cure-pele’, blending occurred in the form of CV/C (/kul/). The same results were found for homophonous words (‘par’ & ‘pare’). For the homophones, blending was performed based on the contextual words. Where the context of homophonous words included words ending in a consonant, blending was performed in the form of C/VC; however, where the homophones occurred in the context of words with final silent ‘<e>’, both forms of blending were observed. The authors explained such performance based on the spelling frequency. Fresh participants tended to spell homophonous words ending in consonant (e.g., pel) rather than with the final silent ‘<e>’. As for pseudowords, the dominant blending form was C/VC. Considering the results of the spelling post-test, participants tended to spell pseudowords ending in consonant rather than in final mute ‘<e>’. Even when fresh participants were asked to write some words sounding similar to pseudowords (so called phonological neighbors), they tended more dominantly to write similar sounding words ending in consonant. Thus, the results of this study suggested that participants relied on the orthographic representations of the neighbor words and blended the syllables of the pseudowords considering the same principles they used for real words. These findings can show that the structure of the syllable of spoken words and pseudowords can also be influenced by orthographic knowledge. Although all pairs had the CVC spoken syllable structure, but processing the internal structure of the spoken syllables was influenced by orthographic information in a purely auditory task.

Orthographic information has also been shown to affect supra-segmental units in the auditory modality. In Thai, tones are lexically distinctive, and they are marked orthographically. In a tone monitoring task, Thai literate participants were asked to judge whether the tones in a pair of spoken words were the same or different (same-different paradigm). The performance was poorer for pairs in which the same tone was represented by different orthographic markers, as compared to pairs in which the same tone was represented by the same orthographic marker (Pattamadilok, Kolinsky, Luksaneeyanawin, & Morais, 2008).

Putting all these results together, it might be argued that the performance in metaphonological tasks is dependent on a strategic application of orthographic representations (Castle & Coltheart, 2004). Literates have two sources of knowledge available in the metalinguistic tasks, phonological representations and orthographic representations, either of which can be used when helpful for task performance. Interestingly, Hulme and colleagues (Hulme, Caravolas, Malkova, & Brigstocke, 2005) showed that orthographic representations are not necessarily used in metaphonological tasks. Two matched[19] groups of children (prereaders) were trained on the initial sounds of words by two different methods. The first group was trained with pictures whose spoken names began with the target phonemes (phoneme group). The second group learned the same target phonemes with corresponding letters (letter group). In both groups the attention was drawn toward the initial sounds of the words. In a third matched group (control group), children received a daily training that was irrelevant to the research purpose. After training, the three groups performed an initial phoneme isolation task in which they were supposed to tell the initial phonemes of the words. The results showed that the phoneme and letter groups outperformed the control group in this task. The authors argue that orthographic representations were not used in the metaphonological task, because if they were, the letter group should have performed better than the phoneme group in the initial phoneme isolation task. This may show that some forms of abstract phonological representations are used in metaphonological tasks.

Considering the findings of the reported experiments with a metaphonological component, it can be claimed that orthographic knowledge helps in creating some form of abstract phonological representations which enable the performance in metaphonological tasks. Such abstract phonological representations can be structured in terms of the orthographic features of a language (Lukatela, Carello, Shankweiler, & Liberman, 1995).

[19] Matched for IQ, reading ability, letter knowledge, and phoneme isolation ability.

2-3. Orthographic knowledge and phonological processing

In the realm of spoken word recognition, the role of orthography has been largely neglected (Pattamadilok, Lafontaine, Morais, & Kolinsky, 2010). However, the evidence shows that orthography and orthographic features may modulate processes involved in spoken word recognition in an *offline* or *online* manner (Taft, 2011). The mechanism underlying the former (*offline* account) assumes that phonological information is restructured in terms of orthographic features. The *online* account, however, assumes that phonological representations coactivate corresponding orthographic representations while processing spoken words. In both accounts, where a phonological unit is represented by multiple spellings and where a phonological unit has no orthographic realization, speech processing is supposed to be hindered. The empirical evidence on this domain shows different outcomes. Some of the empirical evidence has been reported in the following section.

In an auditory lexical decision task, Roux and Bonin (2013) were interested to find whether the number of friends (phonological and orthographic neighbors of a word) could influence spoken word recognition when the feedback consistency[20] was controlled for (Figure 9). They maintained that the influence of orthography on speech processing has been widely investigated by comparing the performance for inconsistent words and consistent words. But, another way to investigate such influence is to consider phonological and orthographic neighbors (friends) of words. The results of the previous experiments showed that words with dense phonological neighborhood were recognized slower than words with sparse phonological neighborhood, but words with dense orthographic neighborhood were recognized faster than words with sparse orthographic neighborhood (Ziegler, Muneaux, & Grainger, 2003). The facilitatory effect of the dense orthographic neighborhood was disappeared when feedback consistency (ratio of friends/enemies) was entered as a covariate. Roux and Bonin assumed that just one type of orthographic neighbors would lead to faster recognition of words, i.e., friends. Feedback consistency could influence recognition latencies in a way

---

[20] Feedback consistency was defined as the ratio of friends by enemies for a given word (Figure 9).

that phonological representations for feedback inconsistent words were not fully specified, and that could lead to slower recognition for inconsistent words as compared to consistent words. To that purpose, they selected matched[21] monosyllabic French words with many or few friends. They were controlled for the body (-VC) consistency. They found that words with many friends (such as B**OULE)** were recognized faster than words with fewer friends (such as B**ARQUE)**. They took the results as an indication for the influence of orthography on the recognition of spoken words by restructuring phonological representations (*offline* account). If the *online* account was the one working while processing speech, the higher the number of friends for an auditory target word, the slower the target word should have been recognized due to the stronger lateral inhibition among the larger number of the friends. However, the results showed the reverse: spoken words with many friends were recognized faster than spoken words with few friends. They reasoned that such a result could be due to the higher resting activation level for words with many friends as a result of developing fully specified phonological representations for consistent words. This finding could show that orthographic knowledge may help creating lexical phonological-orthographic networks (phonographic neighbors) whose members share the same phonological and orthographic units, the higher the number of words in this network, the faster the words can be recognized. Thus, orthography can restructure the organization of the lexical phonological knowledge.

[21] Matched for lexical frequency, number of phonemes and letters, acoustic duration, and uniqueness point.

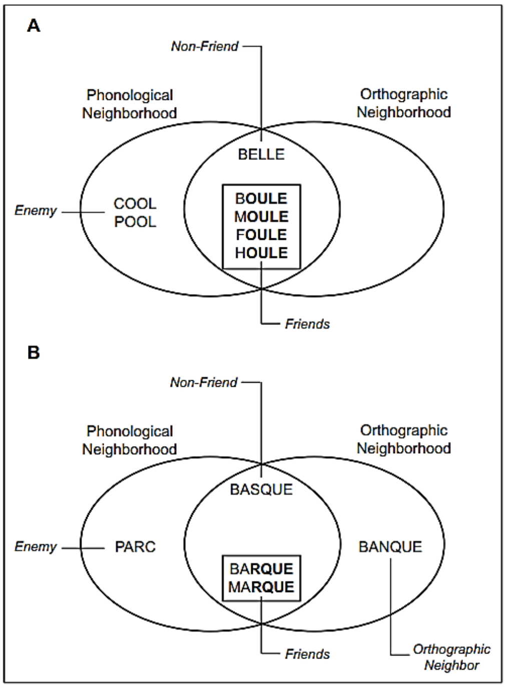


Figure 9. An example of different neighbors for words with many (A) and few (B) friends with the same friend/enemy ratio (Roux & Bonin, 2013), p. 38.

In another auditory lexical decision experiment, the authors (Ziegler & Ferrand, 1998) found that words[22] with consistent rhyme spelling (e.g., /-Ʌk/ in 'buck, tuck, chuck, etc.' in English) rendered faster 'yes' and more correct replies than the words with inconsistent rhyme spelling (e.g., /-i:p/ in 'deep, sheep, heap, leap, etc.'). The authors took the results as an indication for the automatic activation of orthographic information while processing speech (*online* account) in a phonological task that did not require the application of orthographic knowledge.

Later, the results of this experiment were questioned with the argument that they might be due to the inter-stimuli comparison at the decision stage (Ventura, Morais, Pattamadilok, & Kolinsky, 2004). The comparison between the stimuli could have led to rendering spelling knowledge more salient, thus participants recruited phonological as well as orthographic knowledge in making decision on the lexicality of spoken inputs. Lexical decision tasks have been shown to be susceptible to decisional and post

[22] Consistent and inconsistent words were matched for number of rhyme spellings, consistency ratio, frequency, familiarity, number of phonological neighbors, uniqueness point, number of phonemes, and mean duration of spoken words.

perceptual processes that may tap into phonological as well as orthographic information (Ventura, Morais, Pattamadilok, & Kolinsky, 2004). Thus, another study was performed by Ziegler and colleagues (Ziegler, Ferrand, & Montant, 2004) to cope with this critique. To this aim three groups of French words were used as stimuli in three auditory experiments: lexical decision, rhyme detection and shadowing[23] with French adult participants (literates). The stimuli included words with consistent rhyme spelling (e.g., luck, buck, tuck, in English), words with inconsistent dominant rhyme spelling (e.g., wine, nine, fine, line, dine, mine, in English), and words with inconsistent subordinate rhyme spelling (e.g., -ign in sign, design as compared to -ine, in English). In this way, the authors introduced a 'graded consistency' into the study. They designed three different tasks with different levels of difficulty to assess a possible deployment of an orthographic strategy at the decisional stage. Shadowing was the least difficult task with no decisional component. They found the orthographic effect across all tasks: the performance for the inconsistent words with the subordinate spelling was significantly slower than the performance in the two other conditions. Moreover, the performance for the inconsistent words with the dominant spelling was slower than the performance for the consistent words. This effect was strongest in the lexical decision task, moderate in the rhyme detection task, and weakest in the shadowing task. Thus, orthographic information was used in auditory tasks where it was not required or relevant for task performance, and even when the task was easy to perform and did not include a decision component (shadowing). This was interpreted as a robust indication for the automatic application of orthographic information in speech processing.

Ventura and colleagues (Ventura, Morais, & Kolinsky, 2007) conducted two experiments (lexical decision and shadowing) with two groups of French participants, children from grades 2 to 4 and adults, to find whether the results of Ziegler and colleagues (Ziegler, Ferrand, & Montant, 2004) could be obtained. As stimuli, they included consistent (pseudo)words and inconsistent (pseudo)words. If orthographic consistency had an influence on the performance, inconsistent stimuli were expected to have more incorrect replies and slower reaction times as compared to consistent stimuli

---

[23] Immediate repetition of auditory stimuli.

(orthographic consistency effect). In children, Ventura et al. (2007) found an orthographic consistency effect in both tasks for words and pseudowords. However, in adults an orthographic consistency effect was only found in the lexical decision task for words. Thus, a robust orthographic consistency effect was observed for words at the lexical level. The authors interpreted the findings in the frame of *online* account as following. In the early years of learning to read and write, intensive and intentional grapheme-phoneme conversions bring attention to the prelexical level which leads to strong flow of activation at the prelexical level between orthography and phonology (based on BIAM (Grainger, Muneaux, Farioli, & Ziegler, 2005)). However, the sublexical conversions are reduced to a great extent in proficient readers which leads to a weak flow of activation at the prelexical level between orthography and phonology. Thus, consistency effect was not observed for pseudowords in the lexical decision task with adults. On the other hand, shadowing taps into sublexical phonological representations. Reduction in sublexical conversions at the prelexical level can also explain the absence of orthographic effect in the shadowing task with adults. In a 'contingent shadowing' task (repeat stimuli if they are e.g., real words), the orthographic effect was reported (Ventura, Morais, Pattamadilok, & Kolinsky, 2004). The contingent shadowing requires a lexical read-out process ("processes that extract particular representations from memory for specific functionally-based reasons" (Ventura, Morais, & Kolinsky, 2007), p. 569). Their findings thus seem to indicate that the orthographic consistency effect is limited to words and to the lexical level.

Comparing the findings of the studies by the Ziegler group (Ziegler, Ferrand, & Montant, 2004) and the Ventura group (Ventura, Morais, & Kolinsky, 2007), it can be said that the orthographic consistency effect has been observed reliably at the lexical level only. It has not been robustly observed at the sublexical units or prelexical level. Shadowing is a task which does not require lexical access, and pseudowords lack the full lexical representations. For these two cases, no orthographic consistency effect was found in the Ventura study (Ventura, Morais, & Kolinsky, 2007). Considering that both studies were on French, their contradicting results cannot be due to cross linguistic differences. Therefore, to date the question of the exact nature of the orthographic consistency effect in speech recognition remains unresolved.

In addition to the studies in which the influence of orthography was investigated in speech perception or recognition, there are also studies in which the influence of orthography was investigated in speech production. Muneaux and Ziegler (2004) asked participants to produce words that sounded similar to given target words (spoken). The investigators found that participants produced words with similar orthography and phonology more frequently than words with similar phonology but different orthography. For example, when participants were asked to produce a word that sounded similar to 'ripe', they tended to produce 'wipe' more frequently (above the chance level) than 'type'. The results of this experiment suggest an effect of orthographic knowledge in speech production and raise the possibility for the existence of phonographic networks (words with similar phonology-orthography) in the lexicon. Thus, it could show orthographic knowledge may restructure lexicon, so that words with similar phonology-orthography may enter a network created after literacy. Such finding, according to the authors, is in favor of an 'offline' orthographic effect because word frequency played no role in subjects' performances.

The influence of orthography on speech production is not limited to words. Saletta (2019) tested the influence of orthography on the production of pseudowords. The experiment had two phases: a learning phase and a test phase. In the learning phase, participants (poor readers and good readers) heard and repeated pseudowords assigned to pictures, or they read and repeated pseudowords written in opaque (e.g., 'fulvache', /fʌlvæʃ/) or transparent (e.g., 'fulvash', /fʌlvæʃ/) spellings assigned to the same pictures. The segmental accuracy and the articulatory variability[24] of pseudoword production (showing the implicit phonological processing) were measured with a picture naming task after the learning phase. In this task, the participants were asked to name the same target pictures. The segmental accuracy was significantly higher in the group who had read and repeated the pseudowords as compared to the group who had heard and repeated the pseudowords in the learning phase. Moreover, the articulatory variability was significantly higher in the poor readers when producing pseudowords with opaque spelling. Saletta concluded that orthographic information interacted with speech

[24] Articulatory variability was measured by the amount of the movement of lower lip, upper lip, and jaw.

production mechanisms, and that the speech production mechanism was different in readers with different reading proficiency.

Using a picture naming paradigm with written distractors, Lupker and Williams (1988) investigated the effect of orthography on word production. They asked participants to name pictures while ignoring superimposed written distractor words. When the names of target pictures and the written distractors overlapped in sound and spelling (e.g., pl**ane** - C**ane**), the pictures were named faster than when they overlapped in sound only (e.g., pl**ane** - Br**ain**). By contrast, in a study by Damian and Bowers (Damian & Bowers, 2009), where the distractors were presented auditorily, this effect was not found. Comparing the results of these two studies (Damian & Bowers, 2009; Lupker & Williams, 1988), it can be concluded that orthography had an influence on speech production where it was task-relevant, i.e., where it was motivated by the task with written distractors. Therefore, the outcomes could show the strategic rather than automatic application of orthographic knowledge in speech production.

Further evidence for the influence of orthography on speech production comes from a form preparation paradigm originally developed by Meyer (1990). Damian and Bowers (2003) asked participants to memorize English word pairs presented in homogeneous or non-homogeneous conditions. In the former, the word pairs were consistent in phonology and orthography (e.g., **c**amel- **c**offee), whereas in the latter, the words were consistent in phonology but inconsistent in orthography (e.g., '**c**amel-**k**idney'), or inconsistent in phonology and consistent in orthography (e.g., '**c**ycle-**c**obra'). In a test phase following the memorization phase, participants were asked to produce the second word in each pair when cued by the first word. They produced words significantly faster, if they were in the homogeneous condition as compared to the unrelated condition (gypsy- coffee). Such facilitation was not found in the nonhomogeneous conditions as compared to the unrelated condition. This finding could support the influence of orthography on speech production. However, the same results were not obtained in Chinese (Chen, Chen, & Dell, 2002) and Dutch (Roelofs, 2006). Thus, more cross linguistic experiments are required to elucidate the nature of the orthographic consistency effect.

In three different experiments, word reading, object naming, and prompt-response word generation (see the experiment by Damian and Bowers (2003) in the previous paragraph for examples on this paradigm), Roelofs (2006) examined the role of orthography in Dutch word production. He hypothesized that if orthographic information was used mandatorily, the orthographic effect should be observed across all three tasks. However, if using orthographic information was task-dependent, an orthographic effect might be observed in the reading task, but not in the object naming and word generation tasks. The results of his experiments supported the latter. Namely, the response latencies were slowed down significantly by spelling inconsistency in the reading task only. Therefore, he concluded that orthography does not have a mandatory role in word production, but it plays a role in word production only when the task accomplishment requires that. Another picture naming experiment (Alario, Perre, Castle, & Ziegler, 2007) resulted in the same finding. Naming latencies were not significantly different for blocks of partially similar sounding pictures with either homogeneous (consistent) or non-homogeneous (inconsistent) spellings.

In sum, it is difficult to draw a robust conclusion about the nature and role of orthography in speech perception and recognition, as well as speech production considering the conflicting outcomes observed in these series of studies. The topic is still hotly debated, and more empirical evidence is needed.

2-4. Literacy and phonological/cognitive processing

In addition to the studies that compared the performance of literates across orthographically consistent or inconsistent targets, the performance of literates has also been compared with that of illiterates to test the effect of literacy (orthographic knowledge) on speech processing. Such studies show that literacy can lead to developing different strategies in language processing (Petersson, Ingvar, & Reis, 2009).

In an experiment by Ostrosky-Solis et al. (Ostrosky-Solis, Ardila, & Rosselli, 1999), illiterates (zero education) and literates (1 to 3 years of education) were asked to name as many words as possible from the same semantic category, e.g., animal, in a given time (semantic fluency task). In another task, the same participants were asked to name objects beginning with a given unit, e.g., starting with /f/, in a given time (phonological fluency task). The findings showed that semantic fluency and phonological fluency were different across these two groups of participants. In both tasks, the literates outperformed the illiterates (they could name more words). Thus, the literates were phonologically/ semantically more fluent than illiterates. In other words, where the lexicon was to be accessed intentionally by a phonemic/semantic cue, literacy facilitated access to lexicon.

It has also been shown that there is a relationship between literacy and the function of the phonological loop (Silva, Faisca, Ingvar, Petersson, & Reis, 2012). Two groups of adults, literates and illiterates, were studied to find whether they were different in verbal and nonverbal working memory (Silva, Faisca, Ingvar, Petersson, & Reis, 2012). Two tasks were administered: a digit task (verbal) and a spatial task (nonverbal). In both tasks, the stimuli (digits or spatial locations) were to be recalled in forward or backward order. The backward order puts a higher demand on the central executive due to the manipulation of the presented inputs. The findings showed that literacy had a strong influence on performance, in that literates showed a much better performance (more correct replies) in both tasks. The three-way interaction among the factors was also significant. The results of the post-hoc test showed that literates outperformed illiterates in the forward modality of the digit task. Such results could prove that learning an alphabetic writing system could influence the function of the phonological loop. The findings were interpreted by the phonological grain size theory (Goswami & Ziegler, 2005), such that literacy could change the sublexical phonological representations tapped into in the verbal digit task. Literates developed fine-grained sublexical phonological representations that could be retained more efficiently in memory for the recruitment in the verbal tasks.

The same finding was observed in another study with words (instead of digits) by Reis and colleague (Reis & Castro-Caldas, 1997). Literates performed more accurately (84% correct) than illiterates (33% correct) in a pseudoword repetition task (Reis & Castro-Caldas, 1997). In sum, these results suggest that literacy improves the function of the ‘phonological loop’ (Baddeley, 1992), the subcomponent of working memory that maintains phonological information for a limited time.

Later, it was reasoned that the difference in performance between literates and illiterates could be limited to tasks requiring production output. Thus, Petersson and associates (Petersson, Reis, & Ingvar, 2001) decided to test the influence of literacy in a task that didn't require production. In an auditory verbal serial recognition paradigm, the performance of literates and illiterates was compared (Petersson, Reis, & Ingvar, 2001). According to the authors, repetition tasks and verbal serial recognition tasks could entail different cognitive processes, in a way that the immediate serial recognition is not dependent on speech outputs, while the repetition task is. The authors, furthermore, included pseudowords in addition to words to control for the effect of lexicality. The stimuli were either phonologically similar or dissimilar. Two sequences of the same words or pseudowords were presented to literates and illiterates, and they were asked to judge whether the order of the words/ pseudowords was the same in these two sequences (yes/no replies by key press). The main effects of the fixed factors (literacy, phonological similarity, and lexicality) and the interactions among them were significant. The results showed that literates and illiterates were different across all conditions except in the condition with dissimilar words. The results suggested that phonological processing was different in these two groups of participants while lexicality and articulatory or speech output mechanisms were controlled.

Learning an alphabetic writing system can also influence the way pictures are processed and named. The results of an experiment (Reis, Guerreiro, & Castro-Caldas, 1994) comparing the performance of literates and illiterates in picture naming tasks (naming common objects in different presentation modes: drawings, photos, or objects) showed that illiterates were less accurate and slower in naming 2D pictures (drawing presentation mode) as compared to the literate group. These two groups performed the same in 3D picture naming. The authors argue that such a difference may indicate either a difference in processing visual information or in the interaction between processing visual information and the language system. A detailed error analysis was in favor of the former reasoning: there were more visually related errors than language related errors, and there was no reaction time difference between the two groups in naming the real objects. Additionally, a significant positive correlation has been reported between reading level and naming accuracy for line drawings in another study (Goldblum & Matute de Duran, 2000). Thus, the data suggest that literacy has an influence on visual processing: decoding 2D graphemic codes could train literates to perform better in naming 2D pictures.

The findings of experiments with literates and illiterates show that phonological processing is different between these two groups both in tasks that need (picture naming) or do not need (words/pseudoword or digit repetition) lexical access. Literates even showed superior performance in a serial recognition task that does not need any production. Together the findings provide an indication that literates develop a larger short term or working memory span or a higher functionality of the phonological loop. It might be that speech can be better retained in memory when it is additionally represented by a visual orthographic code. The observation that two dimensional pictures were processed more accurately by literates than by illiterates, furthermore, suggests that the effect of literacy goes beyond the language processing system.

2-5. Concluding remarks on the reported literature

Considering the findings of the investigations reported so far, it can be said that they are not converging when it comes to the influence of orthographic knowledge on speech processing. In some of them, orthographic knowledge had no role while processing speech (Damian & Bowers, 2009; Roelofs, 2006; Chen, Chen, & Dell, 2002).Yet others found orthographic effects in speech processing (e.g., (Ziegler, Ferrand, & Montant, 2004; Pattamadilok, Perre, Dufau, & Ziegler, 2009)).

There is also no consensus about the nature of the orthographic consistency effect during speech processing. Some evidence is in favor of the *online* account (e.g., (Pattamadilok, Perre, Dufau, & Ziegler, 2009; Ziegler, Ferrand, & Montant, 2004)); whereas, other findings support the *offline* account (e.g., (Roux & Bonin, 2013)).

Moreover, there is no agreement in the literature whether orthographic effects can emerge at all levels of speech processing (prelexical, lexical, and post lexical). Some studies found orthographic effects at the prelexical level (Taft, Castles, Davis, Lazendic, & Ngye-Hoan, 2008; Ziegler, Ferrand, & Montant, 2004), yet others failed to find it at the same level (Ventura, Morais, & Kolinsky, 2007). With respect to orthographic effects at lexical level, the opposing views of a strategic vs. an automatic influence of orthography are unresolved. Some of the results are in favor of the automatic view (e.g., (Ziegler, Ferrand, & Montant, 2004)), whereas others support the strategic view (e.g., (Damian & Bowers, 2009; Roelofs, 2006; Chen, Chen, & Dell, 2002; Lupker & Williams, 1988), see the introduction in Chapter 7 (Cutler, Treiman, & Van Ooijen, 2010) for an evidence on strategic application of orthographic knowledge in speech processing).

Also, little is known about the nature, the development, and the realization of orthographic effects across languages with different orthographic systems (see e.g., (Pattamadilok, Morais, Ventura, & Kolinsky, 2007) for different size of orthographic effect on speech processing in French and Portuguese). Different mapping patterns are created between phonology and orthography cross linguistically depending on the orthographic transparency. It might be the case that the influence of orthography on speech is related to the nature of such mapping patterns. For example, Taft and associates (Taft, Castles, Davis, Lazendic, & Ngye-Hoan, 2008) measured the response latencies in an auditory lexical decision task for target words whose rhymes could be either orthographically consistent or inconsistent with the rhymes of pseudoword primes. The primes and targets were minimal pairs differing in a vowel. For instance, they primed the target word 'swap' /swɔp/ with the pseudoword /swæp/ in the consistent condition (prime and target were homographs). Or they primed the target word 'stop' /stɔp/ with the pseudoword /stæp/ in the inconsistent condition (prime and target were non-homographs). The relationship between primes and targets was masked from awareness by including many unrelated trials. They found a facilitatory effect in the consistent condition, but not in the inconsistent condition, indicating that the orthography of corresponding primes-targets was active at the sublexical and lexical levels (for a similar design and finding, see (Slowiaczek, Soltano, Wieting, & Bishop, 2003)).

In the language under investigation in that study (English), almost all sounds/phonemes are orthographically represented. What if some vowel phonemes had no orthographic representation like in Persian? If the target word /swɔp/ and the corresponding pseudoword prime /swæp/ were orthographically represented as <swp> instead of <swap>, could the same results be found? One may expect to find different results. In transparent languages such as English as compared to Persian, phoneme-grapheme associations might lead to a more frequent crosstalk between phonology and orthography, and, in consequence, orthographic knowledge may more readily interact with phonological knowledge in speech processing.

Also, due to the frequent cross talk between phonology and orthography in transparent languages, orthographic knowledge can be rendered more salient while processing speech in experiments. Inter-stimuli comparisons may highlight orthographic (in)consistency. By contrast, literates in opaque languages have learned that orthography may not (unambiguously) represent all sounds resulting in a kind of conceptual dissociation between orthography and phonology that can hamper the effect and saliency of orthographic knowledge. In consequence, orthographic effects might be attenuated or not be found at all in opaque orthographies. Such mapping differences may even lead to cross-linguistic differences in the speech processing architecture.

Inversely, the orthographic effect size can be larger in opaque orthographies than in transparent orthographic due to the restructuring of phonological representations in terms of orthographic features which can be stronger in opaque than in transparent orthographies due to higher mismatches or inconsistencies in opaque orthographies.

Consequently, more data are needed to be gathered cross linguistically to confirm whether literates indeed develop a “visual phonology” (Ziegler, Ferrand, & Montant, 2004). Additionally, language processing in illiteracy needs to be studied more intensively. The introduction of spatial representations for temporal representations may open a new window in the minds of literates through which words can be heard or produced in a different way.

Chapter 3: Auditory AX discrimination and phoneme reversal experiments

This chapter has been published as:

Nayernia, L., van de Vijver, R., and Indefrey, P. (2019). The influence of orthography on phonemic knowledge: An experimental investigation on German and Persian. *Journal of Psycholinguistic Research, 48,* 1391–1406. doi:10.1007/s10936-019-09664-9

Note that the numbers of tables, figures, and footnotes have been changed in order to have a consistent numbering throughout this dissertation.

Abstract

This study investigated whether the phonological representation of a word is modulated by its orthographic representation in case of a mismatch between the two representations. Such a mismatch is found in Persian, where short vowels are represented phonemically but not orthographically. Persian adult literates, Persian adult illiterates, and German adult literates were presented with two auditory tasks, an AX- discrimination task and a reversal task. We assumed that if orthographic representations influence phonological representations, Persian literates should perform worse than Persian illiterates or German literates on items with short vowels in these tasks. The results of the discrimination tasks showed that Persian literates and illiterates as well as German literates were approximately equally competent in discriminating short vowels in Persian words and pseudowords. Persian literates did not well discriminate German words containing phonemes that differed only in vowel length. German literates performed relatively poorly in discriminating German homographic words that differed only in vowel length. Persian illiterates were unable to perform the reversal task in Persian. The results of the other two participant groups in the reversal task showed the predicted poorer performance of Persian literates on Persian items containing short vowels compared to items containing long vowels only. German literates did not show this effect in German.

Our results suggest two distinct effects of orthography on phonemic representations: whereas the lack of orthographic representations seems to affect phonemic awareness, homography seems to affect the discriminability of phonemic representations.

Keywords Phonological processing, Phonetic representations, Phonological awareness, Orthography, Grapheme, Orthographical influence, Auditory processing of phonemes, Persian, German

Introduction

By learning to read and write an alphabetic writing system, a connection is made between constituents of written words (letters/graphemes) and the constituents of the spoken forms of words (sounds/phonemes) in memory. A number of researchers have claimed that written constituents can affect the way the spoken constituents are perceived and processed (see e.g., (Kolinsky, 2015; Ehri, 2014). Such effects might not be prominent in transparent orthographies where the grapheme-phoneme mapping across words is consistent. However, opaque orthographies, in which the correspondence between graphemes and phonemes is incomplete or inconsistent, provide a good opportunity to find how a mismatch or inconsistency between written and spoken constituents might influence the perception/production and processing of spoken items.

For example, English speakers hear one more sound in the orally presented words “catch” and “badge” compared to “much” and “page”, respectively (Ehri & Wilce, 1980); and children in fourth grade have problems identifying the phoneme /f/ in ‘laughter’ but not in ‘rafter’ (Castles, Holmes, Neath, & Kinoshita, 2003). In rhyme judgment, words such as “broom-room” were judged faster than those spelled differently (such as “tomb-room”) (Seidenberg & Tanenhaus, 1979). Orthography can even change the number of perceived syllables of a word. For example, speakers who knew the correct spelling of the word “interesting” perceived one more syllable in it than those who spelled it as “*intresting” (Ehri, 1987). Further effects of orthography on explicit phonemic processing have been shown in tasks such as phoneme addition and deletion (Mann & Wimmer, 2002; Caravolas & Bruck, 1993), phonological length judgment (Cassar & Treiman, 1997), and rhyme detection (Prakash, Rekha, Nigam, & Karanth, 1993).

Other studies have found that inconsistency of spelling has an effect on speech perception in an auditory lexical decision task (Pattamadilok, Knierim, Kawabata Duncan, & Devlin, 2010) and in auditory word recognition (Ziegler, Petrova, & Ferrand, 2008; Pattamadilok, Morais, Ventura, & Kolinsky, 2007; Ziegler, Ferrand, & Montant, 2004). For example, spoken words could be recognized faster when they were primed by words that had the same initial phonological and orthographic form:

message-mess. By contrast, words that shared the same initial phonological form, but have different spelling compared to the prime (definite-deaf) were recognized more slowly (Chereau, Gaskell, & Dumay, 2007; Jakimik, Cole, & Rudnicky, 1985).

Orthography can affect word production as well (Saletta, Goffman, & Hogan, 2016; Rastle, McCormick, Bayliss, & Davis, 2011). Saletta et al. (2016), for example, found that orthographic transparency had an effect on auditory pseudoword repetition in adults and children. Non-transparent pseudowords, i.e., pseudowords that could have multiple spellings (e.g., /fispet/ which could be spelled <feespait> or <feespaight>) were produced less accurately, as compared to transparent pseudowords. The repetition ability was measured after a reading task. The accuracy of pseudoword repetitions correlated with the reading level of the participants. The lower the reading proficiency, the less accurate was the repetition of nontransparent pseudowords. The authors concluded that reading and speaking strategies were not the same for proficient and less proficient readers.

Orthography even has an effect on the memory performance for words. In a word form preparation paradigm, Damian and Bowers (2003) found that words from a ‘homogeneous phonology-orthography’ block could be remembered better than words from a ‘homogeneous phonology-heterogeneous orthography’ block, when cued after a memorization phase. As such ‘camel’ was recalled better in a ‘camel-coffee’ block than in a ‘camel-kidney’ block, when cued by the second words. Similarly, Baluch and Danaye-Tousie (2006) found that opaque (non-transparent) Persian words (e.g., <kmk> /komæk/, ’help’) were recalled less accurately as compared to transparent words (e.g., <satur> /sɑtur/, ‘hatchet’).

In contrast to these studies supporting an effect of orthography on speech processing, Ventura and colleagues (Ventura, Morais, & Kolinsky, 2007; Ventura, Morais, Pattamadilok, & Kolinsky, 2004) found no effect of orthography in a shadowing task. These authors, therefore, suggested that there is no effect of orthography on pre-lexical processes and that such effects are limited to lexical processes. There are also authors who claim that the effect of orthography on speech processing may be limited to decision processes or the strategic deployment of orthography (see also (Cutler & Davis, 2012; Cutler, Treiman, & Ooijen, 2010)).

In most studies on effects of orthography on speech processing, polygraphic (a phoneme is represented by more than one grapheme, such as /s/ represented by <c> and <s>), polyphonemic (one grapheme represents more than one phoneme, such as <c> which represents both /s/ and /k/), or silent graphemes (a grapheme has no phonemic correspondence such as <t> in “catch”) have been investigated. In the present study, we targeted the effect of orthographically unrepresented phonemes on speech processing. Orthographically unrepresented phonemes are found in the conventional writing system of Persian. Persian is an Indo-European language with a borrowed Semitic script which replaced Pahlavi script in the year 651 after the islamization of Persia (Hayden, 2018). The conventional written form does not cover short vowel phonemes that are present in the spoken language; hence, the mapping between phonology (phoneme) and orthography (grapheme) can be considered as incomplete in this respect. An unconventional transparent form of the Persian script in which all sounds are orthographically represented is only used in the early months of learning to read and write (6-8 months) after which it is replaced by the conventional form. Persian, therefore, provides a good opportunity to investigate how speech might be processed differently due to long exposure to orthographically unrepresented phonemes in the conventional writing system of Persian.

In the Persian conventional writing system, short vowel phonemes /æ/, /e/, and /o/ are not represented; whereas, the long vowels[25] /ɑ/, /u/, and /i/ have orthographic representations in both the conventional and the unconventional form. For example, <par> /pær/ (feather) in unconventional writing becomes <pr> /pær/ (feather) and is homographic with <pr> /por/ (full) in conventional writing. By contrast, <pir> /pir/ (old) contains one long vowel that is represented in both the conventional and the unconventional forms. Considering the effect of orthography on speech processing, it can, therefore, be hypothesized that items with short vowel phonemes should be processed poorly by monolingual, literate Persian speakers after a long exposure to the conventional writing form.

[25] [1] Letters (among them long vowels) are represented orthographically in different shapes based on their positions in words: in middle and final positions, /i/ is represented as <ﯿ> and <ی> respectively.

We investigated this effect in adult Persian speakers by means of two tasks, an implicit AX discrimination task and an explicit phoneme reversal task. In addition to Persian literate participants, we included two other participant groups for comparison: Persian illiterate adults as a group of speakers who possess phonological but no orthographic representations, and German literate participants (adults) who have been exposed to a script that represents both short and long vowels in its phonology and orthography. Insofar as speech processing is modulated by orthographic features after long exposure to a writing system, we expected that Persian literates (but not Persian illiterates and German literates) to perform significantly more poorly where the processing of items with short vowels is required.

By using two tasks that differed with respect to the levels of processing (see (Morais, 2003) for more information), we aimed to distinguish between possible effects of orthography on the phonemic representations for short vowels at different levels of processing (unconscious and conscious processing). An orthographic influence on the phonemic representations as such would predict a poorer performance of Persian literates in both tasks. By contrast, if orthography is ineffective in prelexical perceptual processing of spoken form (Ventura, Morais, & Kolinsky, 2007; Ventura, Morais, Pattamadilok, & Kolinsky, 2004), we should find a dissociation of performance in the two tasks, i.e., Persian literates should show unreduced performance in AX-discrimination task and reduced performance in the reversal task compared to the control groups.

Experiment 1: AX Phoneme Discrimination Task

Participants

Three groups with 20 participants each took part in the experiment: Persian literates (10 female, mean age 31, age range 20-47), Persian illiterates (14 female, mean age 49, age range 38-60), and German literates (8 female, mean age 30, age range 20-58). For literate participants inclusion criteria were a minimum of 14 years of exposure to written texts and daily reading in their native language. For illiterate participants inclusion criteria were that they had never attended a literacy course and were unable to read simple texts (children’s books) or name the sounds corresponding to Persian letters presented to them in an informal pre-test. All participants were required to have unimpaired hearing/motor response/ speech production.

Persian participants were all from the same region in Iran (Fars province): literates from two major cities, illiterates from a smaller town nearby. German participants were all from the Rhine area. All participants signed an informed consent form and were paid a small sum (8 €) for their participation (i).

Stimuli

We prepared two lists: one contained pairs of German pseudowords or words; the other contained pairs of Persian pseudowords or words. For both languages, pseudowords were created by changing at least one phoneme of a real word and assessed by three native speakers as being possible but non-existing words. Each list was spoken by a female native speaker clearly and fluently in the standard language. The (pseudo)words of a pair were either identical, such as '/korb-korb/' (<Korb>, 'basket') in German, and '/mɑst-mɑst/' (<mast>, 'Yoghurt') in Persian; or non-identical, differing in one vowel phoneme such as '/gold-geld/'(<Gold> 'gold', <Geld> 'money') in German and '/mehr-mohr/' (<mhr>, 'affection', 'stamp') in Persian. Ten warm-ups were included at the beginning of each list.

In eight different conditions (see Table 1 for an overview), non-identical (pseudo)words of a pair differed in one short vowel (S), long vowel (L), or in vowel phonemes with different length (mixed) _ (pseudo) minimal pairs. In the 'mixed' conditions, the (pseudo)words were either homographic (e.g., /xod-xud/ <xud>, 'self/own'- 'helmet') in Persian and /su:xt-sʊxt/ <Sucht-sucht> 'addiction' - (he) 'searches' in German) or heterographic (e.g., as '/bæxt - bɑxt/ <bxt-baxt> 'fate' – 'loss' in Persian and /wo:nən-wɔnən/ <wohnen-Wonnen> 'to live'-'pleasures' in German). Although pseudowords, by definition, do not have any lexicalized orthographic representations, we assumed that literate participants might visualize the written forms of pseudowords while discriminating non-identical pseudo minimal pairs. The pseudo minimal pairs that could have identical orthographic representations (e.g., /ʃantən-ʃa:ntən/ written as <schanten> in German) according to the grapheme-to-phoneme correspondence rules of the respective language were considered as (potentially) ‘homographic’. The homographic pseudo minimal pairs contrast the heterographic pseudo minimal pairs (e.g., /ʃpal-ʃpa:l/ <spall - spal> in German) in the sense that the pairs could have different orthographic forms.

The lists contained 13 non-identical (pseudo)word pairs per condition (in total 104 non-identical pairs) as well as 52 pairs of identical (pseudo)words with long vowels and 52 pairs of identical (pseudo)words with short vowels. The order of (pseudo)words in a pair was counterbalanced across participants. The order of the (pseudo)word pairs in the list was pseudo-randomized across participants with the constraints that (a) no more than two consecutive trials should have the same expected response (‘identical’ or ‘different’) and (b) conditions should not be repeated in consecutive trials [(ii)].

Procedure

The participants were asked to decide by a button press whether two spoken (pseudo)words were identical or different. Each participant was tested alone in a quiet room. A short beep sound signaled a new trial. One hundred milliseconds after the beep sound, a trial consisting of 2 successive spoken (pseudo)words was played through a headset (Sennheiser SC165). The inter-stimulus-interval between the (pseudo)words within a trial was 200 *ms*. After the onset of the second (pseudo)word the participants had 2000 *ms* to press the dedicated buttons for ‘identical’ or ‘different’ responses.

For German and Persian literate participants, the AX phoneme discrimination task session had two parts of 20 min. with a short break in between. In the first part, half of the participants performed the discrimination task in their native language; the other half performed the discrimination task in the unknown language. In the second part the assignment of languages was reversed. Illiterate participants only performed the discrimination task in Persian[26].

[26] [2] Performing all three tasks (Persian and German AX discrimination, Persian reversal task) turned out to be very demanding for illiterate participants, taking in total 2 h including the required training on how to perform the tasks. As most of them became bored and asked to quit the session, we decided to only conduct Persian AX discrimination with these participants.

The AX phoneme discrimination experiment was programmed in Psychopy V.1.82.01 (Peirce, et al., 2019) and run on a Sony laptop. The stimuli were played via a headset and participants pressed the left or right arrow keys (buttons) on the keyboard of the laptop for 'identical' and 'different' responses. The button assignments were reversed for half of the participants.

Results

The proportions of errors (incorrect responses and no responses) per condition are presented in Figs. [1 and 2] 10 and 11 (graphs were generated using SPSS 21, IBM Corp (2012)). We used the lme4 package of R (Bates, Meachler, Bolker, & Walker, 2015) R Core Team (2014)) to perform general linear mixed effect regression analyses with the dependent variable 'accuracy' (correct, incorrect) and the predictor variables 'Literacy' (literate, illiterate) and 'Condition'[27]. As performance on Persian and German tasks could not be meaningfully compared for Persian and German participants, we conducted separate analyses for the two languages. Both predictor variables were entered into the model as fixed effect(s) for Persian participants discriminating Persian minimal pairs. In all other models 'Condition' was the only fixed factor. As random effects, the intercepts for 'participant' and 'word' as well as by-participant random slopes for the effect of Condition were entered into the model. The "multcomp" package (Tukey test) was used to run post-hoc pair-wise comparisons between different levels of the fixed factor(s). *p* values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question.

[27] [3] Levels of Condition: long vowels-pseudowords, short vowels-pseudowords, long vowels-words, short vowels-words, mixed vowels-homographic pseudowords, mixed vowels-pseudowords, mixed vowels-homographic words, mixed vowels-words.

Table 1. Levels of Condition and examples in discrimination task

Condition	Abb.	German				Persian			
		Example	Written	No.of stimuli	Mean no. of phoneme	Example	Written	No. of stimuli	Mean no. of phoneme
Long vowel pseudoword	L-P	/ʃta:n-ʃtu:n/	Stahn-Stuhn	13	4.46	/lik-luk/	لیک-لوک	12	3.17
Short vowel pseudoword	S-P	/broŋ-brɛŋ/	Brong-Breng	13	4.85	/pæŋ-pɛŋ/	پنگ-پنگ	13	3.69
Long vowel real word	L-R	/ʃtu:l-ʃti:l/	Stuhl-Stil chair-style	13	3.77	/fɑl-fil/	فال-فیل fortune telling-elephant	13	3
Short vowel real word	S-R	/hɛft-haft/	Heft-Haft booklet-arrest	13	4.23	/ʃæk-ʃok/	شک-شک doubt-shock	13	3.61
Mixed vowel homograph pseudoword	Mixed-H-P	/ʃantən-ʃa:ntən/	Schanten-schanten	13	5.92	/douf-duf/	دوف-دوف	13	2.92
Mixed vowel pseudoword	Mixed-P	/ʃpa:l-ʃpal/	Spahl-Spall	13	4.46	/pærm-purm/	پرم-پورم	13	3.61
Mixed vowel homograph real word	Mixed-H-R	/ra:stə-rastə/	raste-raste (I)rest-(he)raced	13	4.46	/xod-xud/	خود-خود self/own-helmet	13	2.77
Mixed vowel real word	Mixed-R	/ʃlaf-ʃla:f/	schlaff-Schlaf flaccid- sleep	13	4.38	/bæxt-bɑxt/	بخت-باخت fate-loss	13	3.15

Discrimination in Persian

Persian Participants

As shown in Fig. [1] 10, Persian illiterates made about twice as many errors (error proportion range 0.03 - 0.15) as Persian literates (error proportion range 0.0 - 0.07) across all levels of Condition when discriminating Persian (pseudo)minimal pairs. For both groups, the proportion of errors was the highest when the two stimuli were homographic real words containing vowels of different lengths.

Results of a general linear mixed effects analysis with the dependent variable 'Accuracy' and the predictor variables 'Condition' and 'Literacy' showed that both predictors had significant effects [Condition: $X^2(7) = 18.6$, $p = 0.009$; Literacy: $X^2(1)=15.85$, $p = 0.000$]. The interaction between the two predictors was not significant [$X^2(7)= 5.54$, $p = 0.59$]. Tukey-corrected pair-wise comparisons between the levels of the predictor Condition showed no significant differences (all ps > 0.11).

German Participants

For German literates discriminating Persian word and pseudoword pairs, the proportions of errors ranged from 0.03 to 0.09. Like for the Persian participants, the proportion of errors was highest when the two stimuli were homographic real words containing vowels of different lengths.

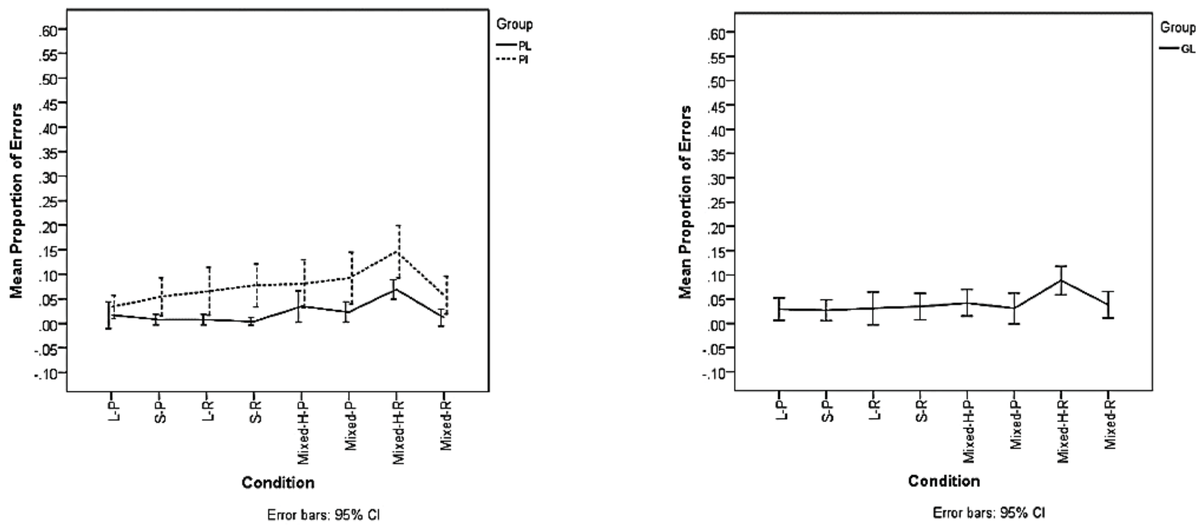


Figure 10. Discrimination in Persian. Left panel: mean proportion of errors for Persian literate (PL) and illiterate (PI) participants. Right panel: mean proportion of errors for German literate participants (GL). (L= long vowels; S= short vowels; H= homographic; P= pseudoword; R= real word).

Results of a general linear mixed effects analysis with the dependent variable 'Accuracy' and the predictor variable 'Condition' showed a marginally significant effect of the predictor Condition [$X^2(7) = 14.022$, $p = 0.051$]. Tukey-corrected pair-wise comparisons between the levels of the predictor Condition showed no significant differences (all *p*s > 0.36).

Discrimination in German

Persian Participants

As shown in Figure [2] 11, Persian literates had no problems discriminating German words or pseudowords containing different long vowels or different short vowels (error proportion range 0.2 – 0.5). By contrast, Persian literates performed much poorer in the 'mixed' conditions, i.e., on German words or pseudowords that differed only in vowel length (error proportion range 0.24 – 0.43).

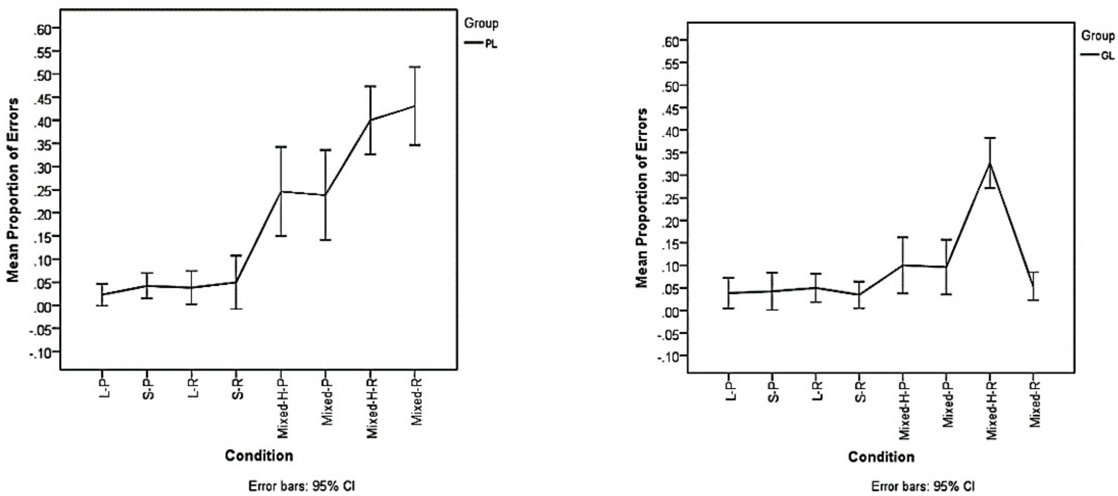


Figure 11. Discrimination in German. Left panel: mean proportion of errors for Persian literate participants (PL). Right panel: mean proportion of errors for German literate participants (GL). (Abbreviations: L = long vowels; S = short vowels; H = homographic; P = pseudoword; R = real word).

Results of a general linear mixed effects analysis with the dependent variable 'Accuracy' and the predictor variable 'Condition' showed that the predictor Condition had a significant effect on the accuracy of Persian literate participants discriminating German minimal pairs [$X^2(7) = 52.59$ $p = 0.000$]. Tukey-corrected pair-wise comparisons between the levels of the predictor Condition showed that the performance of Persian literates on homographic and heterographic real words of mixed vowel lengths was significantly lower than their performance on words and pseudowords with homogeneous vowel length (all ps ≤ 0.05). No other pair-wise comparison reached significance (all ps > 0.07).

German Participants

German literates performed as well as Persian literates discriminating German words or pseudowords containing different long vowels or different short vowels (error proportion range 0.3 – 0.5). The proportion of errors in the mixed conditions was higher (error proportion range 0.05 - 0.33) with a particularly high error rate for homographic real words that differed only in vowel length (0.33). By contrast, the error rate for non-homographic real words that differed only in vowel length was comparatively low (0.05).

Results of a general linear mixed effects analysis with the dependent variable 'Accuracy' and the predictor variable 'Condition' showed that the predictor Condition had a significant effect on the accuracy of German literate participants discriminating German minimal pairs [$X^2(7) = 37.704$, $p = 0.000$]. Tukey-corrected pair-wise comparisons between the levels of the predictor Condition confirmed that the performance on homographic real words that differed only in vowel length was significantly different from all other conditions (all ps < 0.025). No other pair-wise comparison reached significance (all ps > 0.77).

Discussion

Experiment 1 provided no evidence for an influence of the lack of short vowel graphemes on the processing short vowels. Persian literates as well as the two control groups performed equally well on Persian words/ pseudowords containing short vowels as they did on words /pseudowords containing only long vowels. One possibility for the absence of a vowel length effect could be that the task as such did not tap into phonemic representations at all. The participants might have discriminated the input stimuli solely based on their acoustic representations.

A number of findings of Exp. 1 speaks against this possibility. Firstly, Persian illiterates performed significantly worse than Persian literates across all stimulus types. Learning to read and write affects phonemic rather than acoustic representations of words and there is evidence that phoneme boundaries may be sharpened by literacy (for a review see (Kolinsky, 2015)). Hence, the observed difference between the two groups suggests that the participants accessed phonemic representations for discrimination and these representations enabled literate participants to perform better.

Secondly, we observed an unpredicted influence of orthography in German participants performing the same task with German materials. These participants performed significantly worse on homographic word pairs compared to all other conditions suggesting that identical orthographic representations interfere with auditory discrimination. We will come back to this finding in more detail in the general discussion; but as far as the kind of representations is concerned that are accessed in order to perform our discrimination task, it seems more likely that orthographic representations interact with phonemic representations rather than acoustic representations.

Finally, Persian literates had great difficulties discriminating German words and pseudowords that differed only in vowel length (as in <Schote-Schotte>, /ʃo:tə-ʃɔtə/). Such a contrast is not phonemic in Persian (Tsukada, 2011; Campbell & King, 1991) whereas it is phonemic in German (Steinbrink, Groth, Lachmann, & Riecker, 2012) suggesting that the high error rate of Persian listeners is due to the fact that they did not perceive a phonemic difference between these items.

Based on these considerations, we assume that the participants performed our discrimination task by comparing phonemic representations of the input stimuli at least in addition to acoustic representations. The absence of a vowel length effect in Persian literates could, thus, mean that the lack of short vowel graphemes has no effect on phonemic representations. Alternatively, the lack of short vowel graphemes may not affect perceptual phonemic processing, but does affect processes requiring phonemic awareness. In order to distinguish between these two options, we will now turn to Experiment 2 that used a task that required phonemic awareness.

Experiment 2: Reversal Task

Participants

The same participants as in Experiment 1 were asked to participate in the phoneme reversal task in their native language. Despite being trained on the task for at least 30 min., Persian illiterates turned out to be unable to perform the task.

Stimuli

45 disyllabic Persian and 35 disyllabic German words with CVCVC syllable structure were used as stimuli for the Persian and German reversal tasks (see Table 2). The difference in the number of stimuli was due to the shortage of suitable stimuli in German.

The words were selected so that either all sounds in a word were different (C1V1C2V2C3) or the initial and final sounds were identical and the rest were different (C1V1C2V2C1). According to the distribution of the vowel lengths within words, the stimuli fell into different conditions: both vowels of the word (v1 and v2) were short vowels (S), both vowels of the word were long vowels (L), or one of the vowels of the word was short and the other long (mixed). In the 'mixed' condition the two vowels differed both in length and vowel quality. In approximately half of the stimuli of this condition a long vowel was followed by a short vowel. In the other half the order was reversed. All of the words were real words of the respective language and their reversal always resulted in a pseudoword.

The order of the stimuli was pseudo-randomized within lists with the constraint that stimuli from the same condition did not immediately follow each other. Each participant received a different list. All words were spoken clearly and with a normal tempo in the standard form of the respective language by the same speakers as in the discrimination task.

Table 2. Stimuli and conditions in phoneme reversal task

Vowel category	Abbreviation	Persian			German		
Levels		No. of stimuli	Example	Written	No. of stimuli	Example	Written
Short vowel–short vowel	S	11	/pɛsær/	پسر boy	8	/balɛt/	Ballett ballet
Long vowel–long vowel	L	11	/diruz/	دیروز yester-day	6	/t͡suːtaːt/	Zutat ingredient
Short vowel–long vowel	Mixed	23	/muʃæk/	موشک missile	21	/da.tʊm/	Datum date

Procedure

The phoneme reversal experiment was programmed in Psychopy (Peirce, et al., 2019). Each participant was tested in a separate session following a short break after the discrimination task. The experimental task along with some examples was explained to the participants.

An experimental trial started with a beep sound. 200 *ms* later, the spoken stimulus was played through a headset (Sennheiser SC165). From the onset of the stimulus, the participants had 5000 *ms* to reverse the phonemes of the word, and speak the reversed form into the microphone of the headset. In total, the experiment took about 5 *min*. The participants' responses were recorded as audio files on a laptop and analyzed with Praat software (Boersma & Weenink, 2015). The responses were manually transcribed and coded as 'correct' or 'incorrect' by the first author. We considered responses as correct when the two vowels were reversed. Incorrect responses included failures to reverse the vowels, replacements of one or both of the vowels by other vowels, as well as incomplete responses or failures to respond altogether.

Results

The proportions of errors per condition are presented in Fig. [3] 12. General linear mixed effects analyses with the dependent variable 'accuracy' (correct, incorrect) and the predictor variable 'Vowel category' (two long vowels, two short vowels, mixed) were conducted in the same way as for Experiment 1.

Persian Participants

As shown in Fig. [3] 12, Persian literates made more reversal mistakes on Persian words containing at least one short vowel compared to words that contained two long vowels (proportion of errors: two short vowels 0.54; mixed 0.43; two long vowels 0.23). The main error types were vowel changes (46.7 % of all errors), failures to reverse (i.e., reproducing the original stimulus, 24.2 %), and 'no reply' (17.2 %).

Results of a general linear mixed effects analysis (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2014)showed that the predictor Vowel Category had a significant effect on the reversal accuracy of Persian literate participants [$X^2(2) = 12.9$, $p= 0.002$]. Tukey-corrected pair-wise comparisons (Hothorn, Bretz, & Westfall, 2008) between the levels of the predictor Condition confirmed that the performance on words with two long vowels was significantly better than the performance on words containing two short vowels ($Z = 3.360$, $p = 0.002$) and the performance on words containing one long and one short vowel ($Z = 3.679$, $p = 0.001$). The error rates for words that contained two short vowels did not differ significantly from the error rates for words that contained vowels of mixed lengths ($Z = 1.467$, $p = 0.293$).

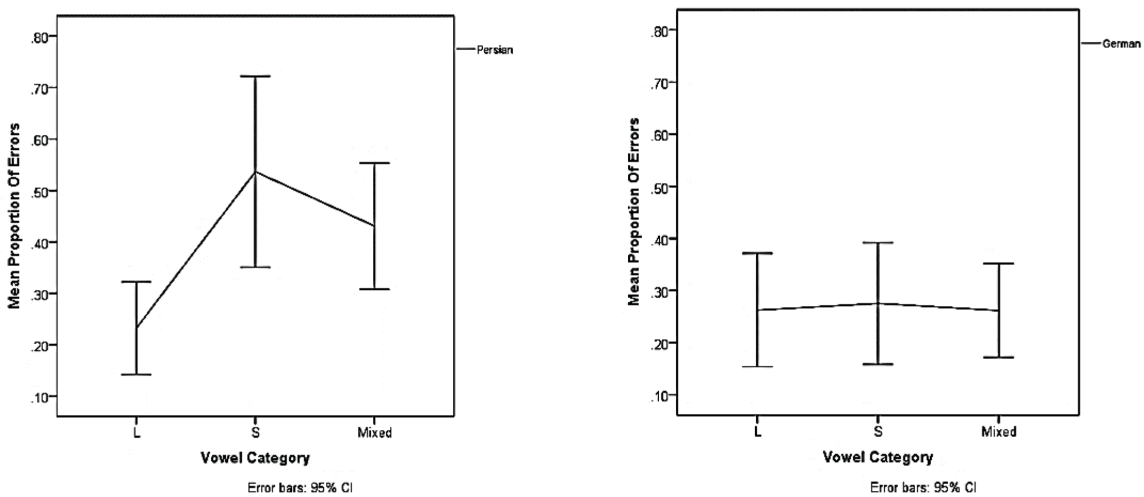


Figure 12. Reversal task. Left panel: mean proportion of errors for Persian literate participants. Right panel: mean proportion of errors for German literate participants. (Abbreviations: L = two long vowels; S = two short vowels; Mixed = one long vowel and one short vowel).

German Participants

For German participants, the mean proportions of reversal errors on German words were similar across the three vowel categories ranging from 0.26 (two long vowels, mixed) to

0.28 (two short vowels). The main error types were 'no reply' (45.5 % of all errors), incomplete responses (23.3%), and failures to reverse (16.3 %). Vowel changes only occurred in 5.4 % of all errors. Results of a general linear mixed effects analysis showed that the predictor Vowel Category had no significant effect on the reversal accuracy of German literate participants [$X^2(2) = 0.004$, $p = 1$].

Discussion

Experiment 2 showed that the presence of a short vowel in the words had a negative impact on the accuracy of Persian literates in the reversal task. By contrast, for German literates the accuracy of the performance was not affected by the vowel category. In the Persian illiterate group, the ability of breaking the integrity of a word into its phonemes could not be attained by a short training session (Morais, Content, Bertelson, Cary, & Kolinsky, 1988). The results thus suggest that the lack of short vowel graphemes affects processes requiring phonemic awareness.

It should be noted that the Persian participants correctly reversed about half of the words that contained one or two short vowels. Furthermore, they never produced phonotactically unacceptable reversed strings (e.g., CCC, CCVC structures). These observations make it unlikely that the Persian participants simply read from the visualized spellings of words containing short vowels.

In vowel change errors, the predominant error type of Persian participants, short vowels were mainly replaced by other short vowels, suggesting that the lack of short vowel graphemes in conventional Persian orthography has a negative effect on Persian literates' ability to identify, maintain, or manipulate short vowels.

General Discussion

In two experiments, we investigated the ability of Persian literate participants to discriminate (Exp.1) and manipulate (Exp.2) words and pseudowords presented auditorily. We hypothesized that due to their long exposure to a script that does not represent short vowels, these participants might show dissociation between their performance on stimuli containing only long vowels and their performance on stimuli containing short vowels. Depending on whether the lack of short vowel graphemes affects phonemic representations underlying speech perception or phonemic representations that can be consciously accessed and manipulated, we expected a poorer performance on stimuli containing short as compared to long vowels in a perceptual task (AX discrimination, Exp.1), in a task requiring phonemic awareness (reversal tasks, Exp. 2) or in both tasks. For comparison, we also presented two other participants groups with the same tasks: a group of Persian illiterates and a group of German literates. We expected none of these groups to show dissociation between their performance on short and long vowels. Persian illiterates had been exposed to spoken Persian to the same degree as the Persian literates; but, they should not show any influence of orthography. German literates had been exposed to written language to the same degree as Persian literates; but, their script does represent both long and short vowels.

The results of our experiments show a clear pattern. There was no evidence for a detrimental influence of the lack of short vowel graphemes on auditory perception; but, there was dissociation between stimuli with and without short vowels in the phonemic awareness task. In Experiment 1, Persian literates discriminated Persian (and German) stimuli containing short and long vowels equally well, and did not differ in this respect from the two control groups. In Experiment 2, Persian literates performed significantly more poorly on the (phoneme) reversal of words containing short vowels than on the reversal of words containing only long vowels; whereas, German literates showed no such difference. In sum, it seems that when phonemes are orthographically unrepresented, they have deficient consciously accessible and manipulable representations; but, the phonemic representations underlying perception are unaffected. This result is in line with Ventura et al. (2004) who assumed two levels for speech processing: a strictly perceptual level and a post-perceptual level. The perceptual level is influenced by linguistic experience and includes modular operations, such as the perceptual segmentation of acoustic representations. At the post-perceptual level, attention and other sources of knowledge affect speech processing as well.

In the Persian script, words that contain only short vowels are homographic. We, therefore, included homographic words also in the German stimuli used in the auditory discrimination task. Most of these stimulus pairs differed only in vowel length. As an unpredicted finding, German listeners performed significantly worse on these stimuli compared to all other stimulus categories, including non-homographic words differing only in vowel length. Persian listeners performed equally poorly on German homographic and non-homographic words differing only in vowel length (numerically even worse on the latter), excluding the possibilities that (a) homographic words just happen to be acoustically more similar than non-homographic words; or (b) that our German speaker inadvertently pronounced homographic words more similarly. It, thus, seems that we observed a true orthographic effect on auditory perception in German listeners: homography has a detrimental effect on auditory discrimination.

This observation raises the question why the lack of short vowel graphemes in Persian also resulting in homography did not have the same effect. It might be the case that Persian literates discriminated the stimuli at a strictly perceptual (pre-lexical) level whereas German literates might have performed the discrimination task at a post-perceptual, lexical level. In other words, depending on language, the same task might have been performed in different ways. According to Pattamadilok et al. (2007), the sensitivity to word spelling might be different for literates in different orthographies. They mention two decisive factors in this regard: first, what type of information (identical/redundant or different) is conveyed by written and spoken forms of a language; second, how robust is the connection between spoken and written forms of a language. In Persian, as compared to German, the information that is conveyed by the spoken form is more comprehensive than that presented by the written form (opaque written form with respect to short vowels). This feature might make Persian listeners rely more on the spoken form. Hence, they might go for pre-lexical strictly perceptual processing for which no orthographic effect is expected. German listeners, on the other hand, might rely more strongly on accessing lexical orthographic representations online, because the German script, at least with respect to vowel length is rather transparent; and hence orthographic representations usually help to distinguish between vowels differing in length.

For the relatively few words where in spite of different vowel lengths the orthographic representations are identical, this strategy backfires as now the orthographic representations signal identity where there is in fact a phonemic difference.

From a computational perspective, for example in the ‘Bimodal interactive activation model’ of Grainger and Ferrand (1996), this situation would correspond to two distinct phonemic representations that both receive feedback from a shared orthographic representation. In consequence, their activation levels become more similar, resulting in poorer discrimination.

In sum, the results of the present study suggest two distinct effects of orthography on phonemic representations: whereas the lack of orthographic representations seems to affect phonemic awareness, homography may affect the discriminability of phonemic representations.

End Notes:

(i) Since the author was considered as an independent researcher, no official organizations supported conducting the experiments. Lack of such support influenced the availability of participants. Most of the people who were invited to participate were unwilling and rejected the author’s invitation, because they assumed that the lack of any official support might cause problems or have negative consequences for them. Persian participants had little or no experience with taking part in a linguistic experiment. So, they might not have been very proficient in focusing on multiple aspects simultaneously, e.g., on the required linguistic knowledge and on task performance. Performing the same experiments with participants from experimental pools might have led to partially different outcomes. The illiterate population was particularly difficult to recruit. Most of them did not want to publicize their illiteracy, and they thought the experiments would make it public to everyone in their small town. The experimenter assured them that no personal information would be leaked out, but they were still reluctant to participate. Illiterates who did participate needed much time and effort to familiarize them with the experimental setting, and due to time constraints they could not be optimally trained. As most of the illiterates were workers who had to work during the day, they participated in the experiments in the evenings or during a short leave from the work. These factors might have affected their performance. Considering all these limitations in sampling, the number of participants was limited to 20-29 participants per experiment. Having a larger sample size would have increased the reliability of the sample mean values.

(ii) The author had to gather the stimuli and their specifications manually due to the lack of a comprehensive free database for spoken/written word frequencies, phonological/ orthographic neighbors, and n-gram frequencies in Persian. Written word frequencies were drawn from the Google search engine, and it had to be assumed that spoken word frequencies correlate with the written word frequencies. Google-reports are well up to date, but Google searches for words in all genres, which cannot be assumed for typical Persian speakers. For instance, not all people read medical or engineering texts.

Chapter 4: Auditory ABX discrimination experiment

4-1. Introduction

According to the Merge model (Norris, 1999), the decision-making mechanism considers all available sources of information (prelexical, lexical, and contextual) while processing speech. The higher the task demand, the higher the sensitivity of the decision-making mechanism to higher-level knowledge. As the auditory ABX discrimination task is more demanding than the AX discrimination task (Morais, Kolinsky, & Claes, 1995), we expected higher level phonological knowledge to be used in the ABX discrimination task as compared to the AX discrimination task. Given that higher-level phonological processing is more susceptible to orthographic knowledge (Pattamadilok, Kolinsky, Ventura, Radeau, & Morais, 2007; Ventura, Morais, & Kolinsky, 2006; Morais, Kolinsky, & Claes, 1995), we reasoned that even if orthographic knowledge did not have a detectable influence in the AX discrimination task (see Chapter 3), it might have an influence at the level of phonological processing, i.e., identification and recognition (Morais, Kolinsky, & Claes, 1995), that the ABX discrimination task taps into.

In the auditory ABX discrimination paradigm, each trial includes three successive spoken items (syllables, words, or pseudowords) in which the third item (X) is identical to the first (A) or the second (B) item in a trial, while A is always different from B (e.g., pat(A)-pet(B), pat(X)). Compared to the AX discrimination paradigm, the ABX paradigm includes a matching component through which the third item in each trial is matched against the first and second items. For this operation, the spoken stimuli should be processed at the recognition-identification level (Morais, Kolinsky, & Claes, 1995). In other words, a similarity comparison must be performed between the items in each trial to arrive at the correct response. According to Treiman and Baron (1981), the similarity comparison can be performed by segmental comparison or by integral comparison (Treiman & Baron, 1981).

Treiman and Baron (1981) assume that for readers, words are composed of segments, whereas such segmental knowledge is missing in prereaders. Prereaders may perceive the stimuli (e.g., words) as an ‘undifferentiated whole’ or ‘integrally’. Gradually, with literacy experience, they learn to analyze the stimuli based on their components. In other words, prereaders are assumed to be ‘holistic perceivers’. The authors illustrate the difference between ‘segmental’ and ‘integral’ similarity comparison with the following example.

In Figure 13 from Treiman and Baron ((1981), p.178), the A and B stimuli have close, but non-identical values on both dimensions, whereas the A and C stimuli have identical values on one dimension, but completely different values on the other dimension. When participants are asked to judge which pair seems more similar (free classification paradigm), they are expected to choose A-B as the more similar pair, if they go for the integral comparison. In this type of comparison, the overall closeness of the stimuli in perceptual space is important rather than the identity on a single dimension. Alternatively, the participants may judge A-C as the more similar pair, because they have an identical value on one dimension. For separable dimensions, the identity on one dimension is very important; and having the identical value on one dimension, it is not anymore important how different the values are on the other dimension.

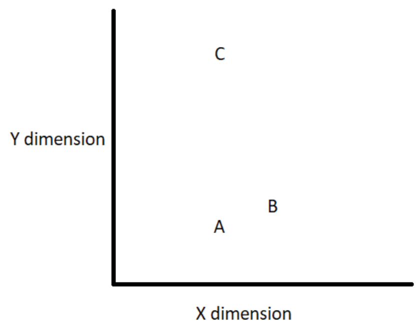


Figure 13. Similarity judgment with three differen stimuli pi (A), te (B), and po (C).

The principles of the similarity comparison proved to be true in visual similarity comparison (Smith & Kemler, 1977). In their experiments, Smith and Kemler asked participants (kindergarteners, second graders, and fifth graders) to categorize stimuli (which two go together?) that varied in size and brightness (two dimensions). The stimuli (a triad in each trial) were created so that they had identical value on one dimension and completely different values on the other dimension, they had close values on both dimensions, or they had completely different values on both dimensions. Three types of classification were possible: participants could put the stimuli in one group that had identical values on one dimension and completely different values on the other dimension (dimensional grouping), they could put the stimuli in one group that had close values on both dimensions (similarity grouping), or they could put the stimuli in one group that had completely different values on both dimensions (haphazard grouping). To find which type of grouping was dominant in each participant group, the mean proportion of grouping in each age was compared against the expected value 0.5. The results showed a significant main effect of age on the type of grouping. That means participants grouped dimensionally with increasing age. Kindergarteners returned significantly less dimensional replies (less than 0.5). Second graders' responses did not differ from the expected value (0.5). Fifth graders' replies were significantly more dimensional (they grouped dimensionally in more than 50% of cases).

Treiman and Baron (1981) tested the principles of similarity comparison in the domain of speech and between spoken stimuli. To prepare the research stimuli, the authors first conducted a similarity judgement. In this task, adults were asked to rate the similarity between different pairs of vowels (such as /i-e/) and different pairs of consonants (such as /p-t/) on a seven-step scale with 1 showing the most similarity and 7 the least similarity. The perceptual similarity of the syllables (CV) was computed by adding the mean rated similarities for the consonants and the vowels in the syllables.

The main task was designed to find an influence of literacy and cognitive development on the similarity comparison of speech units. Two groups of participants (kindergarteners and adult participants) were asked to select the most similar pair in each triad. For example, they were asked to select the most similar pair in a triad like 'pi - po - te'. In this triad, two options were possible: 'pi' and 'po' might be judged as the most similar pair, because they shared the initial consonant (identity in one dimension); thus, the similarity judgment was based on the shared phoneme. Alternatively, 'pi' and 'te' might be judged as the most similar pair because they are perceptually more similar (based on the results of the rating test). In this case, the similarity judgment would be based on the integral similarity of the stimuli (closeness in perceptual space).

The results showed that kindergarteners took the integral similarity of the syllables into account and rated 'pi - te' as the most similar pair, whereas adults tended to base their similarity judgment on the shared phoneme, i.e., they selected 'pi - po' as the most similar pair. Thus, it became clear that depending on age and literacy the principles for the phonological similarity comparison can differ. Cary (1988), furthermore, found that adult illiterates adopted the same strategy as the children in the Treiman and Baron's experiment, suggesting that literacy, rather than cognitive development, was the crucial factor in the change of similarity comparison strategy from integral to segmental.

Morais and colleagues (Morais, Kolinsky, & Claes, 1995) tried to find whether the segments of auditorily presented CV syllables or the integral syllables were attended in the ABX discrimination paradigm. They used two different ABX discrimination tasks (both auditory). In the first ABX discrimination task, the third stimulus (X) was identical to either the first stimulus (A) or the second stimulus (B) in each trial (for ease of reference, it is called the ABX phonological task). In the second ABX discrimination task, the initial phoneme in the third stimulus (X) was identical to the initial phoneme in the first stimulus (A) or to the initial phoneme in the second stimulus (B) in each trial (for ease of reference, it is called the ABX metaphonological task). Two groups of participants, all literates, were asked to decide on the identity of syllables (phonological task) or on the identity of initial phonemes (metaphonological task). The authors expected to find a longer response time in the metaphonological ABX discrimination task as compared to the phonological ABX discrimination task, because in the metaphonological task, the stimuli needed to be segmented and matched for the identity of the initial constituent.

Instead, the authors found equal reaction times in both tasks. Thus, matching initial phonemes and matching the whole syllable took the same time. They concluded that in the phonological as well as the metaphonological ABX discrimination tasks, some sort of covert segmental matching must have been performed. Note that the authors did not test illiterates or prereaders to find whether they would adopt the same strategy in these tasks, because they presumably assumed that illiterates could not perform a task requiring the segmentation of auditory stimuli.

Thus, it can be concluded that higher level phonological processing is involved in the auditory ABX discrimination paradigm as compared to the auditory AX discrimination paradigm. Auditory ABX discrimination requires segmentation of the stimuli and matching the segments against each other to decide whether the third stimulus is identical to the first or the second stimulus in each trial. Hence, the constituents or segments of stimuli must be identified/ recognized, and the phonological knowledge at the recognition level has been shown to be susceptible to orthographic knowledge (Morais, Kolinsky, & Claes, 1995). Moreover, illiterates are not expected to be able to perform the ABX discrimination task based on the segmental strategy, because it requires segmentation and segmental match which only emerge upon literacy (Morais, Kolinsky, & Claes, 1995; Morais, Bertelson, Cary, & Alegria, 1986).

Based on these considerations, we decided to study the effect of orthography on speech processing in an auditory ABX discrimination experiment in Persian. In this experiment, Persian participants were asked to decide whether a third stimulus was identical to the first or the second stimulus in each trial. The trials consisted of triads of Persian real words or pseudowords that either contained long vowels or short vowels. In Persian conventional script, long vowels are orthographically represented, but short vowels are not. If segmentation and segmental match were not influenced by orthographic knowledge, items containing short vowels (e.g., /ʤerm-ʤorm, ʤerm/ <jrm>) were expected to be discriminated as accurately as the items containing long vowels (/bid-bɑd, bid/ <bid-bad>). If, by contrast, segmentation and segmental match were influenced by orthographic knowledge, the items containing short vowels were expected to be discriminated less accurately than items containing long vowels. In other words, homographs (e.g., items with short vowels) were expected to be discriminated less accurately than the non-homographs (e.g., items with long vowels) if orthographic knowledge could influence speech discrimination. An effect of orthography in the ABX discrimination task could be expected if participants adopted a segmental match strategy rather than an integral match strategy.

The same stimuli were used as in the Persian AX discrimination task (Chapter 3), but some modifications were implemented in the ABX discrimination task to reduce the use of acoustic information in the auditory input and to obtain evidence which strategy the participants adopted while performing the ABX discrimination.

First, different persons were used to speak the ‘A-B’ and ‘X’ stimuli in the trials. The A-B stimuli were spoken by the same female speaker as in the AX discrimination task, whereas X was spoken by a male speaker of the same age as the female speaker.

This modification was implemented to reduce the use of acoustic cues to discriminate the stimuli. It has been shown that speaker-specific details can be used to distinguish speech sounds (Smith R., 2015).

Second, two new conditions were added in the ABX discrimination task. These two conditions were mainly intended to provide evidence on the strategy used to perform the discrimination task but would also be sensitive to a possible influence of orthography.

In these two new conditions, A and B differed by an additional vowel (in other conditions, the stimuli within a trial are different by a vowel replacement). For example, in the word pair /?æmud-?æmd/ ‘perpendicular-intention’ a long vowel was inserted, and in the word pair /kerem-kerm/ ‘crème-worm’ a short vowel was inserted. Our reasoning for these conditions was that if participants would use a segmental discrimination strategy without orthographic influence, both additional segments should be readily detectable and lead to good discrimination. If participants would use a segmental discrimination strategy with orthographic influence, long vowel insertion would still lead to good discrimination, but short vowel insertion would not. Because short vowels are not orthographically represented. In trials with short vowel insertion (abbreviated as S-Ø condition), A-B are homographs, /kerem-kerm/ <krm-krm>. On the other hand, if participants use an integral discrimination strategy, the condition with vowel insertion might not lead to better discrimination performance compared to the conditions with vowel replacement. According to Hahn and Bailey (2005), (pseudo)words that are different in a phoneme insertion (/kiθ-kifθ/) sound integrally more similar than (pseudo)words that are different in a phoneme replacement (kif-kiθ/). Thus, if segmental strategy is used to perform the task, the performance across the vowel replacement and vowel insertion conditions should be homogeneous because the items are different in one segment across these conditions. However, if the integral discrimination strategy is used to perform the task, the performance in conditions with vowel replacement should be better than the performance in conditions with vowel insertion, because the items in the vowel insertion condition sound integrally more similar than items in the vowel replacement condition. By adopting an integral discrimination strategy, no influence of orthography was expected since items are compared in the strictly perceptual mode which is assumed to be resistant to the orthographic effect in Persian (see Chapter 3).

As before, we expected an orthographic influence on discrimination performance to be reflected in higher error rates for all conditions with minimal pairs containing short vowels.

4-2. Experiment

4-2-1. Participants

By an announcement through a local person in a small town in the southwest of Iran (Fars Province), about sixty people enrolled to participate in this round of experiments (ABX discrimination with literates and illiterates, phoneme monitoring with literates and illiterates (see Chapter 5), and reading experiments with literates only (see Chapter 6)). As we intended to compare the performance of literate participants across tasks (see Section 6.5), we only used data from literate participants who took part in all experiments. The experiments were administered in different sessions based on the availability of the participants. All participants performed the tasks in the following order: auditory ABX discrimination task, auditory phoneme monitoring task, semantic categorization task with written words, and written pseudoword naming. Persian illiterates were asked to participate in the auditory tasks only.

Twenty-nine literate adults (15 females) and twenty-two illiterate adults (16 females) participated in the experiments of this round including the ABX discrimination experiment. The average age of the literate and illiterate participants was 34.5 (SD=13) and 49.5 (SD=6.5) respectively. All participants were unimpaired in hearing, ability to give motor responses, and speech production. They signed a consent form and were paid a small sum for their participation. Some of the participants in the ABX discrimination experiment had previously participated in the AX discrimination and phoneme reversal tasks (7 literates, 12 illiterates (AX task)).

The literates were native monolingual speakers of Persian who reported daily reading in the Persian script. They were exposed to Persian writing for at least twelve years. The illiterates did not know any letter-sound correspondences in Persian or any other language, and had never been enrolled in a literacy course. In an informal pretest, the illiterates were asked to name Persian letters or read simple Persian words, in which all failed. Both literates and illiterates were trained before the experiment, and they participated in the experiment only after making sure that they had understood the task instructions.

None of the illiterate participants could perform the task in the original time window. This was even the case for those illiterates who had been able to easily perform the AX discrimination task. The illiterate participants tended to repeat the stimuli in each trial several times (some sort of rehearsal) until they felt able to reply. For illiterates, the response time window as well as the inter trial interval were, therefore, extended.

4-2-2. Stimuli

In the ABX discrimination paradigm, three auditory stimuli (A, B, and X) followed each other in a trial where A was always different from B, and X was identical to either A or B (example: bid(A) - bɑd(B), bɑd(X)). The participants were asked to decide by key press whether X sounded identical to A or B. The stimuli in each trial were Persian real words (R) or pseudowords (P).

The experimental stimuli used as the A-B combination in this task were Persian (pseudo)minimal pairs (157 trials) differing in one vowel (critical vowel). The (pseudo)minimal pairs were composed such that either both stimuli contained short vowels (conditions S-S-H-P & S-S-H-R[28]), both stimuli contained long vowels (conditions L-L-P & L-L-R), one stimulus contained a long vowel, the other a short vowel with different spelling (conditions S-L-P & S-L-R), one stimulus contained a long vowel, the other a short vowel with the same spelling (conditions S-L-H-P & S-L-H-R), one stimulus contained an additional short vowel (conditions S- Ø-P & S- Ø-R),

[28] Since short vowels are not spelled, the stimuli in this condition were homograph (H).

or one stimulus contained an additional long vowel (conditions L- Ø-P & L- Ø-R). Table 3 shows the characteristics of the stimuli with examples.

Compared to the Persian AX discrimination paradigm (Chapter 3) there were the following additional modifications in the ABX discrimination paradigm:

First, the identical trials in the AX discrimination task were deleted; since in the ABX task, A is always different from B. Second, 120 consonantal minimal pairs were added as fillers to prevent the participants from focusing on the critical vowels. These consonantal minimal pairs differed in one consonant, and the critical consonant occurred in various positions across trials (examples: initial position /ræd-mæd/; final position /ʃɑl-ʃɑx/). The performance on the consonantal fillers was not analyzed.

In half of the trials within each condition, X was identical to A and in the other half, X was identical to B. The order of A and B stimuli as well as the identity of the X stimulus were counterbalanced across participants. The stimuli were pseudorandomized into 4 different lists and each list was assigned to 7-8 literates and 5-6 illiterates. Within each list, trials from the same condition did not immediately follow each other, and the number of the same expected consecutive responses was limited to three. Ten trials were inserted at the beginning of each list as warm-up trials.

Word frequencies were extracted from the Google engine searching in Persian pages (July 2017, for the words in the new conditions). The reported frequency for each target word was divided by the frequency reported for the most frequent word (/dær/, <dr>, در ‘in, door’) resulting in a relative frequency. The mean relative frequency for the words in each condition is reported in Table 3.

Table 3. Characteristics of the stimuli in the auditory ABX discrimination experiment. Abbreviations: first stimulus in a trial (A), second stimulus in a trial (B), homograph (H), number (No.), Frequency (Freq.), minimal pair (M.P.), number of stimuli per condition is bracketed.

Condition	Abbreviation	Example A	Example B	Spelling A	Spelling B	Mean No. of phoneme A	Mean No. of phoneme B	Mean relative Freq. A	Mean relative Freq. B
M.P differing in a long vowel	L-L-R (13)	bim	bɑm	BIM 'fear'	BAM 'roof'	3	-	0.026	0.034
Pseudo M.P differing in a long vowel	L-L-P (12)	lik	luk	LIK	LUK	3.17	-	0	0
M.P differing in a short vowel (H)	S-S-H-R (13)	ʤerm	ʤorm	JRM 'mass'	JRM 'crime'	3.61	-	0.036	0.036
Pseudo M.P differing in a short vowel (H)	S-S-H-P (13)	pæŋ	peŋ	PNG	PNG	3.69	-	0	0
M.P differing in a vowel from different categories	S-L-R (13)	bænd	bɑnd	BND 'string'	BAND 'band'	3.15	-	0.09	0.056
Pseudo M.P differing in a vowel from different categories	S-L-P (13)	zæm	zɑm	ZM	ZM	3.61	-	0	0
M.P differing in a vowel from different categories (H)	S-L-H-R (13)	xod	xud	XUD 'self'	XUD 'helmet'	2.77	-	0.122	0.122
Pseudo M.P differing in a vowel from different categories (H)	S-L-H-P (12)	touz	tuz	TUZ	TUZ	2.92	-	0	0
M.P with a short vowel Insert ion (H)	S-Ø-H-R (16)	kerem	kerm	KRM 'crème'	KRM 'worm'	4.4	4.62	0.056	0.056
Pseudo M.P with a short vowel Insertion (H)	S-Ø-H-P (13)	belz	belez	BLZ	BLZ	4.4	4.7	0	0
M.P with a long vowel Insertion	L-Ø-R (13)	?æmud	?æmd	?MUD 'perpen-dicular'	?MD 'intention'	4.5	4.7	0.029	0.056
Pseudo M.P with a long vowel Insertion	L-Ø-P (13)	sædf	sædɑf	SDF	SDAF	4.6	4.8	0	0

4-2-3. Procedure

Each participant was tested individually. The experiment started with the task instruction. When participants were ready, they pressed the start button to begin the experiment. 300 *ms* after pressing the start button, a calm beep sound was played to signal the new trial. 100 *ms* after the signal, the first stimulus in each trial (A) was spoken. 250 *ms* after the end of the first stimulus, the second stimulus (B) in that trial was spoken. 500 *ms* after the end of the second stimulus, the third one (X) was spoken.

From the onset of the third stimulus in each trial, the participants had 1000 *ms* time to press the response button on the keyboard of the laptop. They were instructed to press the left arrow key, if X (the third stimuli in each trial) was identical to A (the first stimulus in each trial), or to press the right arrow key if the X was identical to B (the second stimulus in each trial). The position of the response key was changed for half of the participants. The inter trial interval was set to 1500 *ms*. At the end of the experiment, a 'thank you' note was played. The experiment was run on a Sony laptop using Psychopy V1.82.01 (Peirce, et al., 2019). The participants listened to the stimuli wearing a headphone (Sennheiser SC165).

Literate participants could perform the task in this time frame; however, illiterates could not. Thus, the timing was increased in steps until they might perform the task. Four of the illiterates could do the task when the interval between A and B was 250 *ms* (no change in the standard timing), the interval between B and X was set on 1000 *ms* (500 *ms* longer than the standard timing), and the response time was set on 5000 *ms* (4000 *ms* longer than the standard timing). However, according to their reports, they did not press the response keys systematically in most of the cases even then.

4-3. Data analysis

As described above, even with extended time windows, only few illiterates could perform the ABX task. Also, according to their feedback in the debriefing session, they pressed the response keys by guessing most of the time. A preliminary analysis of their responses also confirmed that. Therefore, the data from illiterate participants had to be excluded from further analysis.

In the literate group, the reaction time for 15% of the experimental trials was between 0 and 100 *ms* (mean reaction time for correct replies to all experimental trials: M= 715 *ms*, SD=293). These extremely fast replies were interpreted as the mis-manipulations of the response button and not considered for further analysis. 8.3% and 33.7% of the literates' replies were incorrect and no response trials respectively. The high proportion of no response trials suggested that when participants were not completely sure of the correct response, they refrained from responding. This was confirmed by their reports in the debriefing session. For the statistical analysis incorrect and missing responses were collapsed as 'errors'. Figure 14 (generated by SPSS 21 (IBM Corp., 2012)) shows the proportion of errors as a function of 'vowel category' for real word and pseudoword pairs.

General linear mixed effect regression models (using lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2014)) were used to assess the influence of the factors 'vowel category' (S-S, L-L, S-L, S-L-H, S- Ø, and L- Ø) and 'lexicality' (real word vs. pseudoword) on the accuracy (correct vs. incorrect (error)) of discrimination in the auditory ABX discrimination task. The intercepts and slope of 'items' and 'participants' were entered into the model as random factors. The main effect of each factor was computed by comparing the model including all factors against the model without that target factor (likelihood ratio test - ANOVA). Post hoc pairwise comparisons between the levels of the factors were performed using the multcomp package – Tukey test (Hothorn, Bretz, & Westfall, 2008) where the corrected p-values were reported based on the number of comparisons.

4-4. Results

The results of the Persian ABX discrimination task (Figure 14) for real words showed that minimal pairs differing in a short vowel (S-S-H-R: M=0.28, SD=0.038) or differing in a long vowel (L-L-R: M=0.26, SD=0.034) were discriminated with lower proportions of error as compared to minimal pairs differing in a vowel from different categories (S-L-R: M=0.37, SD=0.03 and S-L-H-R: M=0.52, SD=0.04). The error rates were the highest when participants discriminated minimal pairs with a vowel insertion (S-Ø-R and L-Ø-R: M=0.55, SD=0.03).

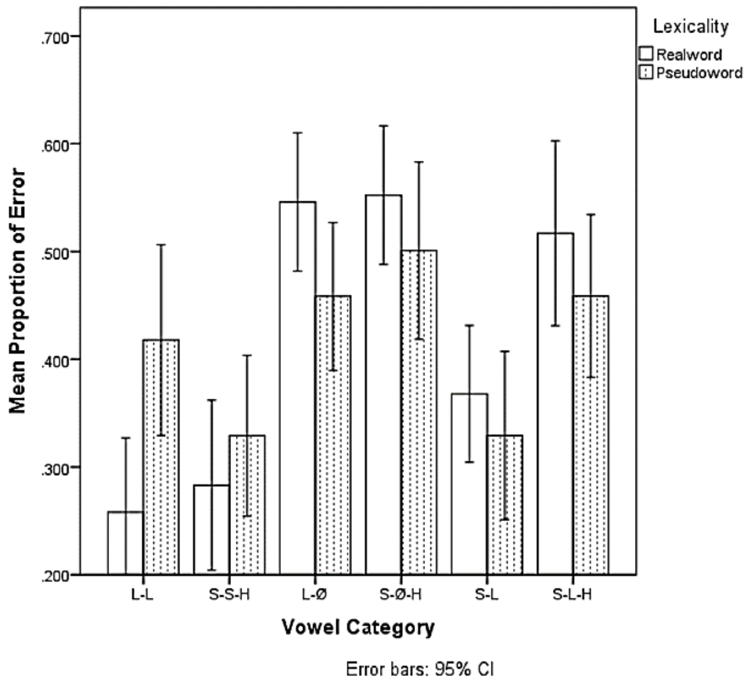


Figure 14. Auditory discrimination (ABX) of Persian (pseudo)words by Persian literates. Mean proportion of errors as the function of vowel category and lexicality. Abbreviations: Short vowel (S), Long vowel (L), missing the critical vowel (Ø), Homograph (H).

For the discrimination of pseudoword pairs, the error rates were the lowest for non-homograph pseudoword pairs differing in vowels from different categories or homograph pseudoword pairs differing in a short vowel (S-L-P and S-S-H-P: *M*=0.33, *SD*=0.04). Pseudoword pairs differing in a long vowel (L-L-P: *M*=0.42, *SD*=0.04), homograph pseudoword pairs differing in vowels from different categories (S-L-H-P: *M*=0.46, *SD*=0.04), and pseudoword pairs with vowel insertion (L-Ø-P: *M*=0.46, *SD*=0.033 and S-Ø-P: *M*=0.5, *SD*=0.04) were discriminated with higher error rates.

Considering only homograph pairs, homograph minimal pairs differing in a short vowel were discriminated with a lower proportion of errors than homograph minimal pairs differing in vowels from different categories, or homograph minimal pairs with short vowel insertion.

The results of the general linear mixed effect regression analysis with ‘lexicality’ and ‘vowel category’ as the independent variables and ‘accuracy’ of discrimination as the dependent variable showed a significant main effect of ‘vowel category’ [$X^2(10)$ =47.028, p<0.001], but no significant main effect of ‘lexicality’

[$X^2(6)=10.15$, $p=0.12$]. The interaction between these two factors was marginally insignificant [$X^2(5) = 10.148$, $p=0.07$].

The detailed pairwise comparisons between the levels of 'vowel category' for word pairs[29] showed significant differences between the discrimination of homograph and non-homograph minimal pairs differing in a vowel from different categories ($Z=2.95$, $p=0.02$), between the discrimination of minimal pairs with short vowel replacement and short vowel insertion ($Z=-4.26$, $p<0.001$), and between the discrimination of minimal pairs with long vowel replacement and long vowel insertion ($Z=4.74$, $p<0.001$). Also, the discrimination was significantly different between minimal pairs with long vowel replacement and homograph minimal pairs differing in a vowel from different categories ($Z=3.91$, $p=0.001$) or minimal pairs with a short vowel insertion ($Z=4.7$, $p<0.001$). The pairwise comparison between the discrimination of minimal pairs with long vowel and short vowel replacement did not turn a significant value ($Z=0.46$, $p=1$).

Pairwise comparisons between the levels of vowel category for pseudoword pairs showed statistically no significant difference between the performance across the levels (all $ps>0.1$).

The results of a general linear mixed effect regression model for the effect of homography (with the two levels 'homographs' and 'non-homographs') on discrimination accuracy showed no significant effect of homography [$X^2(1)=1.46$, $p=0.23$].

4-5. Discussion

Compared to the AX discrimination task, the ABX discrimination task was more demanding, because based on the literature (see section 4-1 (Morais, Kolinsky, & Claes, 1995)) it included a segmental matching component in addition to the discrimination component. Hence, it was assumed that orthographic knowledge could influence the

[29] The reported results are limited to the corresponding levels relevant to the study purpose.

performance in the auditory ABX discrimination task. We, therefore, expected that Persian literate participants would discriminate minimal pairs containing short vowels less accurately than minimal pairs containing long vowels. In the Persian conventional script, long vowels are orthographically represented, but short vowels are not. Hence, minimal pairs containing short vowels are homographs, and, to the degree that orthographic knowledge had an influence, this property should interfere with discrimination in the ABX paradigm. Assuming that participants would use a segmental matching strategy, we, furthermore, predicted that discrimination would be particularly easy in two conditions where the number of segments was different due to the insertion of an additional vowel in one of the (pseudo) minimal pairs within a trial.

The results of the auditory Persian ABX discrimination task, however, were complex and did not confirm our predictions. Although there were overall significant differences between conditions, the pattern of differences was not compatible with our prediction that short vowels should be more difficult to discriminate. The discrimination accuracy was the same when participants discriminated minimal pairs differing in a short or in a long vowel replacement (/sæm-som/ <sm> or /bid-bɑd/ <bid-bad>) as well as when participants discriminated non-homograph minimal pairs containing vowels from different categories (/bænd-bɑnd/ <bnd-band>). Comparatively, the discrimination accuracy was significantly lower when participants discriminated minimal pairs with a short or long vowel insertion (/kerm-kerem/ <krm> or /?æmd-?æmud/ <?md-?mud>) as well as when participants discriminated homograph minimal pairs differing in a vowel from different categories (/xod-xud/ <xud>). Thus, there was no clear evidence that the lack of orthographic representations for short vowels or homography was led to poor discrimination in the ABX discrimination task.

In other words, the overall findings show that (i) minimal pairs with short vowel replacements were discriminated as accurately as minimal pairs with long vowel replacements, but less accurately than minimal pairs with short vowel insertions, (ii) minimal pairs with long vowel replacements were discriminated more accurately than minimal pairs with long vowel insertions, and (iii) homograph minimal pairs differing in a vowel from different categories were discriminated less accurately than non-homograph minimal pairs differing in a vowel from different categories or minimal pairs differing in a long or short vowel replacements.

The only finding that might be seen as reflecting some effect of homography was the higher error rate of homograph minimal pairs differing in a vowel from different categories (/xod-xud/ <xud>) compared to non-homograph minimal pairs differing in a vowel from different categories (/bænd-bɑnd/ <bnd-band>). However, given that other conditions did not show an effect of homography and that a separate regression analysis with homography as a factor did not show a significant effect of this factor, an alternative explanation for the higher error rate in this condition seems much more likely. Note that, as explained earlier, the homography of the items in this condition (homograph minimal pairs differing in a vowel from different categories /xod-xud/ <xud>) is considered as a form of 'irregularity' in Persian by which the short vowel /o/ comes to be written with the same grapheme as the long vowel /u/ in a small number of exception words. In consequence, the vowel pair /o-u/ was to be discriminated in the homograph mixed condition while in the non-homograph mixed condition, the vowel pair /æ-ɑ/ was to be discriminated. According to the related literature (Hahn & Bailey, 2005), similarity judgment was a function of the number of matched sub-phonemic features between two words. The higher the number of mismatched sub-phonemic features, the less similar the two words sound. According to this point, the discrimination accuracy was expected to be the same for /xod-xud/ and /bænd-bɑnd/ because the vowels in each pair are different in one sub-phonemic feature which is the relevant factor in sound discrimination according to Hahn and Bailey. In /o-u/, the vowels are [back, mid - back, high]; and in /æ-ɑ/, the vowels are [front, low-back, low]. But the results of the ABX discrimination task showed that the vowel pair /o-u/ was discriminated less accurately than the vowel pair /æ-ɑ/; thus, it seems that the number of different sub-phonemic features is not the relevant factor in discrimination in the ABX discrimination task. The different error rates of the two conditions may be due to the vowel pair /u/-/o/ being perceptually closer than other vowel pairs. Indeed, a study on the perceptual distances between Persian vowels (Rahmani, Bijankhan, & Peyvasteh, 2007) reports the following ranking:

(u-e) > (ɑ-e) > (o-e) > (u-i) > (i-e) > (e-a) > (o-ɑ) > (ɑ-a) > (u-o)

Given that the perceptual distance between /u/ and /o/ is the smallest of all vowel pairs, it seems that the relatively high discrimination error rate in the homograph mixed condition relying on this contrast is best explained by perceptual distance playing a significant role for the judgments in the ABX discrimination task (see (Morais, Kolinsky, & Claes, 1995) for similar findings on poor discrimination for perceptually more similar stimuli as compared to perceptually less similar stimuli).

Apart from the homograph mixed condition, error rates were significantly higher in the two conditions with vowel insertion (/ʔæmd-ʔæmud/ <ʔmd-ʔmud> and /kerem-kerm/ <krm-krm>) compared to the conditions involving short or long vowel replacements. This result is the opposite of what we predicted based on the assumption that participants would use a segmental matching strategy in the ABX discrimination task. Coming back to the distinctions presented in the introduction section (Section 4-1) between the integral and segmental matching strategies (Treiman & Baron, 1981) as well between phoneme insertion and phoneme replacement (Hahn & Bailey, 2005), it seems that participants in our experiment did not use a segmental strategy, but somehow took the overall form of the auditory stimuli into account, such that vowel replacements were better discriminated or perceived as less similar than vowel insertions were.

Hahn and Bailey (2005) found that a mismatch between aligned components affected the perceived integral dissimilarity more than a mismatch between non-aligned elements. Hence, /kiθ/ as a target was perceived as more similar to /kifθ/ than to /kif/.

The performance in the ABX task also suggests that participants adopted predominantly an integral strategy rather than a segmental strategy to judge about the identity of the stimuli in a trial. Although the minimal pairs across all experimental conditions were different in one vowel phoneme, some of them could be discriminated better than others. The minimal pairs with vowel insertion (e.g., /kerem-kerm/ in S-Ø-R condition or /ʔæm**u**d-ʔæmd/ in L- Ø -R condition) sounded integrally more similar than the minimal pairs with a vowel replacement (e.g., /ʤerm-ʤorm/ in S-S-R or /b**i**d-b**ɑ**d/ in L-L-R); hence, the discrimination was lower for the former (vowel insertion) as compared to the latter (vowel replacement).

Thus, the findings of the ABX discrimination experiment suggest that discrimination was performed using an integral matching strategy (Treiman & Baron, 1981). Adopting an integral matching strategy might be due to the fact that the response decisions could be achieved more easily and faster given the time constraints in the experiment.

Another reason that could have led to adopting an integral match strategy was the task design. The minimal pairs were structurally aligned in 222 trials of the ABX task (including fillers), whereas only 55 trials comprised items that were not structurally aligned (L-Ø and S-Ø conditions). The higher number of structurally aligned trials might have biased our participants to adopt an integral matching strategy. This strategy could lead easily to the correct reply since the differences are more salient in structurally aligned stimuli (Hahn & Bailey, 2005). For structurally not-aligned stimuli, a segmental match was necessary for the correct reply.

Alternatively, it might be the case that Persian literates are holistic perceivers[30] as a result of the orthographic depth of the Persian conventional script. It has been shown that the coupling grain size between orthography and phonology is larger in opaque orthographies than in transparent orthographies (Goswami & Ziegler, 2005). Reading in an opaque script with orthographically unrepresented short vowels may have trained Persian literates to process units larger than segments in the auditory modality, too. Thus, it remains to investigate whether participants becoming literate in a transparent orthography would perform the ABX discrimination task with (pseudo)word items by adopting a segmental match strategy.

The results of our experiment differ from the results of the experiment by Morais and colleagues (Morais, Kolinsky, & Claes, 1995) who found that literates adopted a segmental strategy in the ABX discrimination task. Such difference can be interpreted considering the materials used across these two experiments. The stimuli in the present experiment were composed of words and pseudowords that were created by

[30] As explained in the introduction section to this chapter, integral match or closeness in perceptual space is the crucial factor for holistic perceivers (Treiman & Baron, 1981).

changing at least one sound of the real words. The stimuli in the experiment by Morais group were nonsense CV syllables. Using words and pseudowords, instead of nonsense syllables, as stimuli may have encouraged participants to adopt an integral matching strategy.

In contrast to the AX discrimination task, illiterate participants were unable to perform the ABX discrimination task. Given that error rates were also considerably higher for literate participants, the ABX discrimination task is obviously more demanding. Whether illiterates failed, because the task itself was cognitively too demanding or because they were unable to perform the required matching operations is unclear, and further experiments would be needed to decide between the two possibilities.

In sum, the results of the auditory ABX discrimination task with Persian literates did not show any influence of orthography. Vowels with and without orthographic representations were discriminated with the same accuracy. However, the integral perceptual similarity between the (pseudo)words in minimal pairs did influence the discrimination accuracy, such that the higher the integral perceptual similarity, the lower the discrimination performance. The lack of an orthographic effect on the performance in the ABX discrimination task can be explained by the level of processing at which the participants approached the task. In contrast to the prediction that the participants would go for a segmental matching strategy allowing for the deployment of higher level phonological knowledge i.e., identification-recognition (Morais, Kolinsky, & Claes, 1995), the results rather suggest that the ABX discrimination task was performed at a perceptual level, and an integral rather than a segmental matching strategy was used to perform the task. At the perceptual level, most previous experiments also did not find an orthographic effect (Ventura, Morais, & Kolinsky, 2007; Pattamadilok, Kolinsky, Ventura, Radeau, & Morais, 2007; Morais, Kolinsky, & Claes, 1995), compatible with the suggestion that speech processing is immune to orthographic knowledge at the perceptual level (Karmiloff-Smith, 1995).

Chapter 5: Auditory phoneme monitoring experiment

5-1. Introduction

In the phoneme monitoring task (alternatively called phoneme identification/ detection), a target phoneme is to be detected in carrier units (e.g., syllables, words, pseudowords). There are three proposals on how the target phoneme can be detected.

According to the first proposal, a target phoneme can be identified after the recognition of the word in which the target phoneme is embedded (Marslen-Wilson, 1984; Foss & Swinney, 1973). Thus, the lexical phonological codes stored in the lexicon are used to detect the target phoneme. In fact according to this proposal, the phoneme is a unit that can be accessed post-lexically. The second proposal states that the target phoneme can be identified based on the phonetic information extracted from the auditory input signal (Foss & Gernsbacher, 1983); hence, the lexicon is not involved. According to the third proposal (autonomous model), the information from both sources can be used to detect a target phoneme (Foss & Blank, 1980; Cutler & Norris, 1979): a phoneme can be identified based on sublexical and/or lexical codes. The code (sublexical or lexical) that becomes available first is used to detect the target phoneme.

The position of target phonemes in words plays a role with respect to the involvement of prelexical or lexical information in phoneme detection tasks. The identification of initial phonemes is mainly performed based on prelexical information (Frauenfelder & Segui, 1989). Also, when the experimental list includes just monosyllabic items in which the target phonemes are to be detected, the chance of the involvement of lexical information is reduced (Cutler, Mehler, Norris, & Segui, 1987). The findings of phoneme detection experiments show that phoneme detection is faster:

- when the target phoneme occurs in the initial position of accented syllables (Pitt & Samuel, 1990; Shields, McHugh, & Martin, 1974).
- in words than in nonwords (Frauenfelder, Segui, & Dijkstra, 1990; Cutler, Mehler, Norris, & Segui, 1987). However, no word advantage was found in an experiment by Foss and Blank (1980).

- in natural speech as compared to synthesized speech, i.e., error rates are lower when detection is to be done on normal speech (Nix, Mehta, Dye, & Cutler, 1993).

It has, also, been shown that the rate of missed responses is negatively correlated with vowel duration when detecting English vowels in normal speech (Cutler, Van Ooijen, Norris, & Sanchez-Casas, 1996): the longer the vowel duration, the lower the missed response rates. This inverse correlation was not found between vowel duration and missed response rate in a Spanish phoneme detection experiment with Spanish participants (Cutler, Van Ooijen, Norris, & Sanchez-Casas, 1996). The authors interpreted the differential findings of the English and Spanish experiments in terms of the variability of the vowel system in these two languages; the Spanish vowel inventory being sparser and well distinct compared to English. Moreover, length is not phonemic in Spanish. Hence, different factors may play a role in detecting vowels cross linguistically depending on the linguistic properties of languages.

Moreover, it has been claimed that phoneme perception is different from phoneme identification which is a component of phoneme detection (Swinney & Prather, 1980; Foss & Swinney, 1973). In phoneme identification (included in phoneme monitoring), conscious access to phonemes is required, whereas in phoneme perception, unconscious representations of phonemes would be sufficient. Foss and Swinney (1973) explained the processes involved in phoneme detection as follows: first acoustic-phonetic features are extracted from the input signal (perception of the auditory input), then the larger units (words, syllables) are identified. At last, the conscious representation of the phoneme is accessed.

A number of experiments have investigated the effect of literacy and orthographic knowledge on the performance in phoneme monitoring. In this line of research, the experiment presented in this chapter investigated whether phonemes with no orthographic representations can be detected as accurately as phonemes with orthographic representations, and whether illiterates can detect phones/phonemes embedded in carrier units. I will first report the findings of some of the experiments in which the effect of orthography and literacy on phoneme monitoring has been studied.

In a phoneme detection experiment (Van Ooijen, Cutler, & Norris, 1992), the literate participants were asked to press a response key as soon as they detected a target phoneme anywhere in spoken word stimuli. Although in this auditory phoneme detection task, the focus was expected to be on the spoken form of the stimuli rather than on their spelling, the results of the experiment showed that the number of missed responses was significantly higher in words such as ‘dune, cubic, fuse’ compared to words such as ‘yard, yolk, yell’, when /j/ was to be detected. Thus, when the target phonemes had no canonical orthographic realization, their detection was disrupted. This could be an indication that orthographic representations could facilitate the detection of a target phoneme (for a similar finding and interpretation in the case of schwa detection in English, see (Cutler, Norris, & Van Ooijen, 1990)).

In another study, Frauenfelder and colleagues (Frauenfelder, Segui, & Dijkstra, 1990) found that /k/ was detected more slowly than /p/ or /t/ in French words. They attributed this difference to spelling consistency. In French, /k/ is inconsistently spelled in multiple ways <c>, <k>, or <q(u)>, whereas, /t/ and /p/ are consistently represented in writing. An effect of orthographic consistency on phoneme monitoring has also been reported in Dutch (Dijkstra, Roelofs, & Fieuws, 1995). Furthermore, orthographic consistency effects have been reported not only in phoneme monitoring but also in syllable (Dupoux, 1993) and rhyme (Donnenwerth-Nolan, Tanenhaus, & Seidenberg, 1981) monitoring. All these findings suggest that either the coactivated orthographic representations while processing speech have influenced phoneme detection performance, or the phonemic representations have been restructured in terms of orthographic features.

In another experiment on the sensitivity of phoneme detection to orthographic knowledge, Cutler and colleagues (Cutler, Treiman, & Van Ooijen, 2010) tried to gain more insight in the nature of the orthographic effects in auditory phoneme monitoring. They investigated how /b, m, t, f, s, k/ phonemes were detected by English literates in different experimental contexts. The first three phonemes of this set /b, m, t/, in contrast to the last three ones /f, s, k/, are consistently spelled in English. The critical phonemes occurred word-initially in English words. They found that the inconsistently spelled phonemes /f, s, k/ were detected more slowly than the consistently spelled phonemes /b, m, t/. However, this difference was only observed when the list of stimuli contained many words with irregular spellings as fillers, e.g., ‘kneel, wrinkle, tongue, etc.’, in which some letters are not pronounced. Inconsistently spelled phonemes did not show slower detection latencies when the list contained filler words with regular spellings, e.g., ‘nip, wiggle, lung, etc.’ Therefore, the authors suggest that orthographic knowledge was strategically used in phoneme detection when spelling inconsistency was rendered salient in the experimental context by including many irregularly spelled English words.

To investigate the effect of literacy on the phoneme detection ability, Morais and colleagues (Morais, Bertelson, Cary, & Alegria, 1986) asked Portuguese illiterates and ex-illiterates to monitor for the following predefined targets: the stressed syllable [tá], the phone [r], the unstressed syllable [ta], and the phone [k]. The critical units occurred in word-initial position in five sentences. The participants were expected to find the words that contained the targets. The accuracy of the performance was analyzed. The ex-illiterates outperformed the illiterates. Across both participant groups, the stressed syllable [tá] was detected better than the unstressed syllable [ta]. Likewise, [r] was detected better than [k] in both groups of participants ([r] is produced with a great energy in Portuguese). In the illiterate group, the detection score for [k] was much lower than that for syllables. Also, [k] was detected less accurately by illiterates than by ex-illiterates, whereas the accuracy of the syllable detection was almost identical across the participant groups. Subsequently, a rhyme detection task was assigned to the same group of participants to test their sensitivity to the rhyme unit as opposed to phones and syllable units. In contrast to the illiterates’ poor performance in detecting phones, illiterates performed as well as ex-illiterates in detecting rhymes, suggesting that sensitivity to rhymes and syllables can be attained in the absence of literacy, whereas the detection of phones is literacy dependent. Moreover, the accessibility of phones (e.g., /r/ in Portuguese) could facilitate phone detection.

In sum, it can be concluded that effects of orthography have been found in auditory phoneme detection experiments that do not require orthographic information to be recruited. There is, to date, no agreement if orthographic codes are activated strategically or automatically in phoneme detection. Additionally, the findings suggest that conscious representations of phonemes are used in phoneme detection tasks (Foss

& Swinney, 1973) as illiterates have problems in detecting phonemes, and literacy plays an important role in creating conscious representations of phonemes (Morais, 2021; Goswami, 2000).

In the following experiment, we investigated whether phonemes with orthographic representations (long vowels) are detected as accurately as phonemes without orthographic representations (short vowels) in Persian words and pseudowords. In case orthographic representations were coactivated in addition to phonological representations in phoneme detection or in case phonemic representations were restructured in terms of orthographic features, we expected the phonemes /æ, e, o/ (short vowels) to be detected less accurately than /ɑ, u, i/ (long vowels) due to the lack of any orthographic representations for the former, but not for the latter.

5-2. Experiment

5-2-1. Participants

Persian participants (literates and illiterates) who had taken part in the ABX discrimination task (Chapter 4) were asked to also participate in the phoneme monitoring experiment. However, in spite of receiving a short training before the experiment, Persian illiterates were unable to perform the task.

5-2-2. Stimuli

In Persian, there are 6 vowels /æ, e, o, ɑ, u, i/, from which the first three are short vowels /æ, e, o/ without orthographic realization in the conventional writing. The other three are long vowels /ɑ, u, i/ with orthographic realizations in the conventional writing. Six different lists were prepared. In each list, one of the vowels was to be detected (target vowel).

Each list included disyllabic words and pseudowords (30 items each) carrying the target vowel, and the same number of disyllabic fillers (60 items: 30 real words, 30 pseudowords) that did not carry the target vowel. For example, the list for detecting the vowel /æ/ included words and pseudowords such as /kælɑɣ/, <klagh>, 'crow' and

/pætɑʃ/, <ptash>, and fillers such as /livɑn/, <livan>, ‘mug’ and /puken/, <pukn>. The pseudowords were created by changing at least one phoneme of Persian real words.

The target vowels occurred in second position after /p/, /t/ or /k/ in the first syllable of the words and pseudowords. 60% of the fillers had also /p/, /t/ or /k/ as the initial phoneme to prevent predicting the intended vowel by the initial consonant of each item. The first syllable that contained the target vowel was unstressed.

The target vowels were either in the medial position or in the offset of the first syllable. For example, the target phoneme /æ/ was in the offset of the first syllable in /k**æ**-lɑɣ/ C**V**-CVC, and in the medial position of the first syllable in /t**æ**ʃ-viɣ/ C**V**C-CVC. Half of the experimental items carried the target vowel in the offset position of the first syllable, whereas in the other half, the target vowel occurred in the medial position of the first syllables.

The (pseudo)words were randomized in three different orders in three lists, and each list was assigned to 1/3 of participants. Within each list, the (pseudo)words were pseudorandomized in such a way that at most three consecutive trials had the same expected response (‘yes’ or ‘no’). Ten items were used as warm-ups at the beginning of each list. Table 4 represents the characteristics of the stimuli in the phoneme monitoring experiment.

The word frequencies were extracted from the Google search engine (December 2017). The relative frequencies were computed by dividing the reported frequency for each carrier word by the frequency of the most frequent word /dær/, <dr> ‘in, door’. The mean relative frequency for the words in each condition is reported in Table 4.

Table 4. Characteristics of the stimuli in the phoneme monitoring experiment. With examples (Ex.), and the relative mean frequency (freq.) of stimuli (real word (R) and pseudoword (P)).

Vowel	**æ (short)**		**e (short)**		**o (short)**		**ɑ (long)**		**u (long)**		**i (long)**	
lexicality	**R**	**P**	**R**	**P**	**R**	**P**	**R**	**P**	**R**	**P**	**R**	**P**
Ex.	/kælaɣ/	/pærɑl/	/kelid/	/pemut/	/tohmæt/	/pomder/	/kɑseh/	/pɑmut/	/pulæk/	/kupsɑn/	/kiseh/	/tinkes/
Meaning	crow		key		insult		bowl		spangle		sack	
Spelling	KLAGH	PRAL	KLID	PMUT	THMT	PMDR	KASH	PAMUT	PULK	KUPSAN	KISH	TINKS
Mean relative freq.	0.0109		0.0088		0.0082		0.0107		0.0087		0.0092	

5-2-3. Procedure

In the phoneme monitoring task, each participant was tested individually. They were asked to listen to the spoken (pseudo)words that were played successively and press the respective response key on the laptop keyboard as soon as they detected/ did not detect the target vowel in the spoken items. If the target vowel was detected in the spoken items, they were instructed to press the key labeled 'yes'; otherwise, they were instructed to press the key labeled 'no' (the translation of 'yes/no' in Persian was used). For half of the participants, the position of the response keys was changed. The experiment was run on Psychopy software (Peirce, et al., 2019) installed on a Sony laptop. The participants listened to the spoken stimuli through a Sennheiser SC165 headphone.

The experiment had 6 parts; in each part, one of the vowels was to be detected. At the beginning of each part, the participants were informed which vowel was to be detected by an instruction. Afterwards, the participants pressed the spacebar for the experiment to begin. By pressing the spacebar, a beep sound was played which signaled a new trial. 200 *ms* after the beep sound, the spoken item was played. The participants were to decide within 2000 *ms* from the onset of the spoken item, whether the target vowel was / was not in the spoken item and press the corresponding response key accordingly. Then, the next trials were presented one by one. The inter trial interval was set on 1000 *ms*. At the end, a 'thank you' note indicated the end of that part of the experiment. Each part was followed by a short pause. The responses were recorded in an excel output sheet.

5-3. Data analysis

None of the Persian illiterates was able to detect the target vowels in the spoken items. Even after a short training session before the experiment and extending the original response time stated above for the illiterates, they failed to perform the task. As a result, data could only be collected from the Persian literates.

7% and 13% of the experimental items were followed by no-replies and incorrect replies. Missing and incorrect responses were collapsed as 'errors' for the statistical analysis. The graphs below show the detection error rates as a function of 'vowel category' and 'lexicality' (Figure 15) and as a function of 'vowel category', 'vowel duration', and 'lexicality' (Figure 16).

General linear mixed effect regression and Tukey tests were used to analyze the data and perform pairwise comparisons (lme4 (Bates, Mächler, Bolker, & Walker, 2015) & multcomp (Hothorn, Bretz, & Westfall, 2008) packages). Two models were used to analyze the accuracy of the detection data: in the first model, the accuracy of the replies (correct vs. incorrect (error)) was analyzed as a function of the predictors 'vowel category' (short vowel vs. long vowel) and 'lexicality' (real word vs. pseudoword). In the second model, the accuracy of the replies was analyzed as a function of the predictors 'vowel category', 'vowel duration' (long, mid, short), and 'lexicality'. Note that in contrast to the levels 'long', 'mid', and 'short' of the predictor 'vowel duration', the labels 'short vowel' and 'long vowel' for the levels of the predictor 'vowel category' refer to an orthographic distinction in the modern Persian script, but not to the actual duration of the vowels. The second model was used because the literature showed an effect of vowel duration on the detection accuracy for vowels in English (Cutler, Van Ooijen, Norris, & Sanchez-Casas, 1996). 'Participants' and 'items' were inserted into all models as random factors.

The effect of each factor was calculated by ANOVA tests comparing the model including all factors with the model without that target factor. The corrected *p*-values for the number of comparisons are reported in the text.

As noted earlier, the position of the target vowels was varied within the syllable in this experiment. The target vowels were either in the medial position or in the offset position of the first syllable of the spoken items. A preliminary analysis showed that the accuracy of the performance was statistically the same across the medial and offset positions [$X^2(1) = 0.75$, $p=0.36$]; thus, the data were collapsed for the position of the vowel within the syllable.

5-4. Results

5-4-1. Accuracy as a function of 'vowel category' and 'lexicality'

As shown in Figure 15, Persian literates detected short vowels (S) with a higher mean proportion of errors as compared to long vowels (L) in both real words (M_S=0.27, SD=0.21; M_L=0.12, SD=0.15) and pseudowords (M_S=0.28, SD=0.2; M_L=0.15, SD=0.14).

The results of the general linear mixed effect regression analysis for the effect of 'vowel category' (long vowel with orthographic representation vs. short vowel without orthographic representation) and 'lexicality' (real word vs. pseudoword) on the accuracy of the vowel detection showed a significant main effect of 'vowel category' [$X^2(2)$ =80.9, p<0.0001] on the accuracy of the vowel detection. 'Lexicality' had no significant effect on the accuracy of vowel detection [$X^2(2)$ =3.29, p=0.19]. Also, the interaction between 'vowel category' and 'lexicality' was insignificant [$X^2(1)$ =0.93, p=0.34].

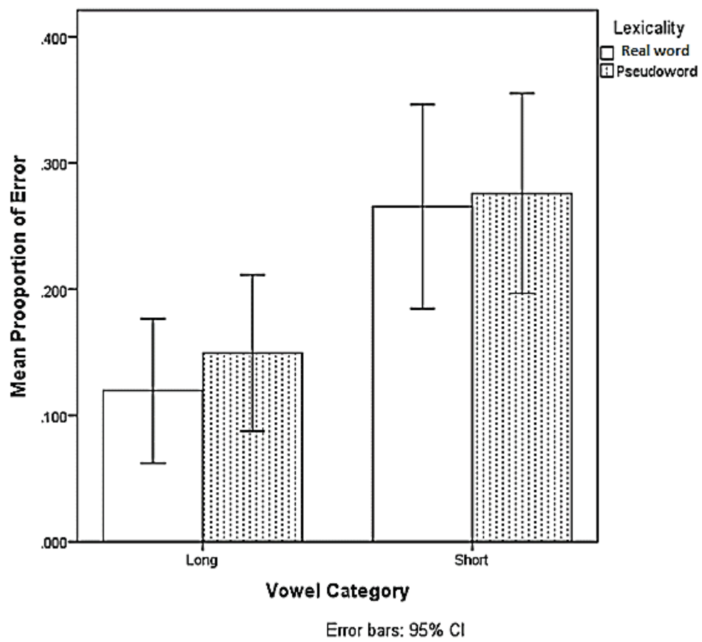


Figure 15. Auditory phoneme monitoring by Persian literates. Mean proportion of errors as a function of vowel category and lexicality.

5-4-2. Accuracy as a function of ‘vowel category’, ‘lexicality’, and ‘vowel duration’

Since the literature showed that the duration of vowels could influence the vowel detection rate in English (Cutler, Van Ooijen, Norris, & Sanchez-Casas, 1996), a separate analysis was performed to find how ‘vowel duration’ could influence the performance in the phoneme monitoring task. In this analysis, ‘vowel category’, ‘lexicality’, and ‘vowel duration’ (high, mid, and short) were considered as the fixed factors; and their effects on the accuracy (correct vs. incorrect (error)) of the detection were analyzed.

According to Sheikh Sangtajan and colleague (Sheikh Sangtajan & Bijankhan, 2013), the average durations of Persian vowels are /o/=280 *ms*; /æ/=239 *ms*; /e/=249 *ms*; /ɑ/=275 *ms;* /u/=252 *ms*; /i/=230 *ms*. Based on these values, Persian vowels were grouped into three categories: ‘long duration’ (/o/, /ɑ/), ‘mid duration’ (/u/, /e/), and ‘short duration’ (/æ/, /i/). In the following section, the results of the analyses considering vowel duration are reported.

Figure 16 shows the mean proportion of errors as a function of ‘vowel category’ and ‘vowel duration’ for real words and pseudowords. In the real word panel, the

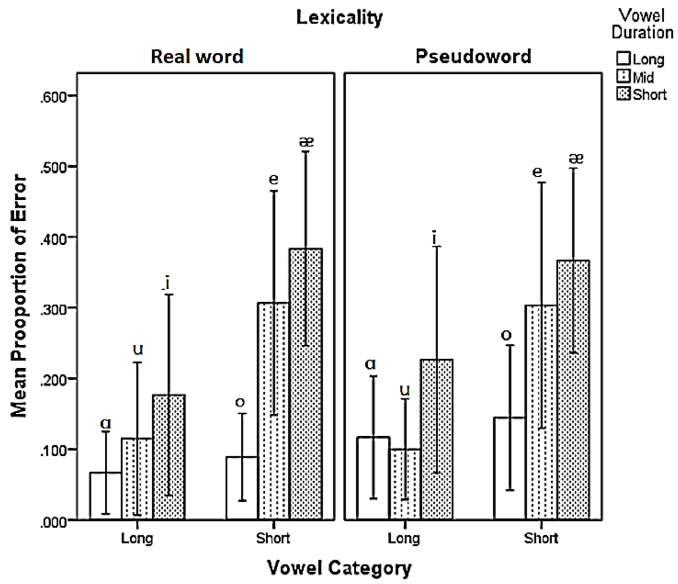


Figure 16. Auditory phoneme monitoring by Persian literates. Mean proportion of errors as a function of vowel category and vowel duration for real words and pseudowords.

predictor 'vowel duration' correlates inversely with the proportion of errors; that is, the longer the vowel duration, the lower the mean proportion of errors. An almost similar pattern is observed in the pseudoword panel. Across 'vowel category', vowels without orthographic representations were detected with higher proportion of errors (real word: *M*(æ) =0.38, *SD*=0.3, *M*(e)=0.31, *SD*=0.2; pseudoword: *M*(æ)=0.37, *SD*=0.18, *M*(e)=0.3, *SD*=0.24) than the vowels with orthographic representation (real word: *M*(u)=0.11, *SD*=0.14, *M*(i)=0.18, *SD*=0.2; pseudoword: *M*(u)=0.1, *SD*=0.09, *M*(i)=0.23, *SD*=0.22) when vowels had mid or short durations. Vowels with (real word: *M*(ɑ) = 0.07, *SD*=0.08; pseudoword: *M*(ɑ) =0.12, *SD*=0.12) and without (real word: *M*(o) =0.09, *SD*=0.08; pseudoword: *M*(o) =0.14, *SD*=0.14) orthographic representations were detected with approximately the same proportion of errors when they had long duration.

The results of the general linear mixed effect regression analysis showed that 'vowel category' (short vs. long) and 'vowel duration' (short vs. mid vs. long) had significant effects on the detection accuracy ['vowel category': $X^2(6) = 105$, $p<0.0001$; 'vowel duration': $X^2(8) = 135$, $p<0.0001$]. However, 'lexicality' had a marginally insignificant effect on the accuracy of the detection [$X^2(6) = 11.5$, $p=0.08$].

The results of the pairwise comparisons between the levels of 'vowel duration' indicated that the vowels with long duration were detected more accurately than the vowels with mid ($Z=-6.46$, $p<0.0001$) or short ($Z=-7.7$, $p<0.0001$) duration; but the vowels with mid duration were detected as accurately as the vowels with short duration ($Z=-0.62$, $p=0.8$).

As for the three- and two-way interactions among and between the fixed factors and their effects on the accuracy of the detection, the following results were obtained. The interaction among 'vowel category', 'vowel duration', and 'lexicality' was significant [$X^2(2) =25$, $p<0.0001$]. The interaction between 'vowel category' and 'vowel duration' was also significant [$X^2(2) = 16$, $p<0.001$]. The two-way interactions between 'vowel category' and 'lexicality' [$X^2(1) = 0.9$, $p=0.35$] as well as between 'vowel duration' and 'lexicality' [$X^2(2) = 5.8$, $p=0.08$] were insignificant.

The results of the detailed analyses between the levels of the predicators 'vowel category', 'vowel duration', and 'lexicality' showed that[31] when participants detected the vowels with long duration (the short vowel (o) and long vowel (ɑ)), the error rates were significantly lower in real words than in pseudowords (Z(ɑ)=2.1, p=0.036; Z(o)=2, p=0.046). Other pairwise comparisons between pseudowords and real words containing the short vowel with mid duration (Z(e) =-0.09, p=0.93), containing the short vowel with short duration (Z(æ) =-0.42, p=0.67), containing the long vowel with mid duration (Z(u) =-0.56, p=0.58), or containing the long vowel with short duration (Z(i) =1.52, p=0.13) did not show significant values. Furthermore, when participants detected vowels with long duration, there was no significant difference between the detection accuracies for short and long vowels in real words (Z(o-ɑ)=-0.99, p=0.32) and pseudowords (Z(o-ɑ)=-0.98, p=0.33); but, there were significant differences between the detection accuracies for short and long vowels in real words and pseudowords, when participants detected vowels with mid (real word: Z(u-e)=-5.4, p<0.0001; pseudoword: Z(u-e)=-5.72, p<0.0001) or short (real word: Z(i-æ)=-5.53, p<0.0001; pseudoword: Z(i-æ)=-3.73, p=0.0002) duration.

The results of the detailed analyses between the levels of the predictors 'vowel category' and 'vowel duration' showed that the short vowel with long duration was detected as accurately as the long vowel with long duration (Z(o-ɑ) =-1.38, p=0.17). But, the short vowels with mid or short duration were detected less accurately than the long vowels with mid (Z(u-e) =-7.85, p<0.0001) or short (Z(i-æ) =-6.55, p<0.0001) duration.

5-5. Discussion

The results of the auditory phoneme monitoring experiment with Persian literates showed the following results. (i) Orthographically unrepresented vowels were generally detected less accurately than the vowels with orthographic representations. (ii) Vowels of long duration were detected more accurately than vowels of short or mid duration, but vowels of mid duration were detected as accurately as vowels with short duration.

[31]The reported results are limited to the corresponding levels relevant to the study purpose.

(iii) Importantly, we observed a significant interaction between 'vowel category' and 'vowel duration': The effect of orthography (vowel category) was not detectable for vowels of long duration, but vowels of mid or short duration were detected significantly better if they were orthographically represented.

The first results showed that in a completely auditory task where no recourse to orthographic information was required, orthographic knowledge was used. This is in line with what Van Ooijen and colleagues (Van Ooijen, Cutler, & Norris, 1992) found: the phonemes with orthographic realization (such as /j/ in 'yell') were detected more accurately than the phonemes without any orthographic representation (such as /j/ in 'dune').

The second finding confirms the results of an experiment by Cutler and associates (Cutler, Van Ooijen, Norris, & Sanchez-Casas, 1996). In an English vowel detection task, these authors found that vowel duration correlated inversely with the proportion of missing responses, suggesting that the identification of vowels with short duration was more demanding than identification of the vowels with long duration.

The third finding could show the nature of the orthographic effect. The significant interaction between 'vowel category' and 'vowel duration' suggests that the presence of an orthographic effect can depend on a phonetic property. Such performance can be accounted for by the Merge model for spoken word recognition (Norris, 1999). This model assumes a decision-making mechanism that considers all available sources of information to make a decision about whether the auditory input contains a target phoneme. When the input provides a rich phonetic information in the case of vowels with long duration, the mechanism may put more weight on the phonetic information rather than orthographic information. By contrast, in case of vowels with mid or short duration, the orthographic information may be weighted more strongly in the process of decision making. As a result, an orthographic effect was found for vowels with mid or short duration, but not for vowels with long duration.

Alternatively based on the results of Cutler et al. (Cutler, Van Ooijen, Norris, & Sanchez-Casas, 1996), the interaction we observed might be due to the relative timing of the identification of vowels with short or long durations. For vowels of long duration, the availability of rich phonetic information might lead to the fast correct detection of this vowel, and hence the orthographic effect might not have found time to emerge. Conversely, for vowels of mid or short duration, the delay in detection might have allowed the orthographic effect to emerge. There is evidence in the literature suggesting that orthographic effects may emerge late in speech processing. Pattamadilok group (Pattamadilok, Morais, & Kolinsky, 2011) observed an orthographic consistency effect in a shadowing task when auditory stimuli were presented in noise. However, in a 'no noise' condition, no orthographic consistency effect was found in the same study. The finding suggests that the noise-induced delay in the processing of the auditory stimuli, was necessary for an influence of orthographic knowledge to emerge. The recruitment of orthographic knowledge, thus, must be relatively late.

With respect to the illiterates, it can be argued that they failed to perform the task, because the detection of phonemes requires the segmentation of the spoken inputs and matching the segments against a sample (the target phoneme). The illiterates failed to break the integrity of the auditory inputs. Phoneme monitoring tasks have a metaphonological component, and performing them needs access to the conscious representations of phonemes that are developed after literacy or similar training (Morais, 2021; Morais & Kolinsky, 1995; Morais, Kolinsky, & Claes, 1995). Thus, illiterates could neither perform the phoneme reversal (Chapter 3) nor the phoneme monitoring tasks in this study.

By comparing the results of the Persian phoneme reversal (Chapter 3) and phoneme monitoring tasks, three points become clear. First, Persian illiterates failed to perform both tasks. Second, the performance for long vowels was better than the performance for short vowels in both tasks in the Persian literate group. Third, the pattern of the performance varied for the individual vowels across these two tasks in the Persian literate group.

In the reversal task, the performance for every orthographically unrepresented vowel was less accurate than the performance for every orthographically represented vowel. In the phoneme detection task, the detection of the orthographically unrepresented vowel /o/ was as accurate as the detection of the orthographically represented vowel /ɑ/ and more accurate than the detection of other orthographically unrepresented vowels. These different patterns of performance in the phoneme reversal task and the phoneme monitoring task can be interpreted in terms of the level of processing (Karmiloff-Smith, 1995). The two tasks seem to tap into different explicit representational formats. The phoneme monitoring task may rely on an explicit representational format that is available to consciousness, but not manipulable or available to verbal reflection. By contrast, the phoneme reversal tasks may rely on another explicit representational format that is manipulable and available to consciousness and verbal reflection.

In other words, the representational formats used in these two tasks might be different in the level of abstraction. Phoneme reversal might require the manipulable format of representation, i.e., the highly abstract representation of phonemes (called E3 format by Karmiloff-Smith), whereas phoneme monitoring could be performed on a less abstract representational format (called E1/E2 format by Karmiloff-Smith). Hence metalinguistic tasks may vary with respect to the representational format they use.

To conclude, the findings of the phoneme monitoring experiment showed that orthographic knowledge was used in the auditory phoneme monitoring task in Persian. The orthographic knowledge interacted with the phonetic information (vowel duration) suggesting the strategic deployment of orthographic knowledge. The findings demonstrated that where the vowels were of the mid or short duration, those represented orthographically were detected more accurately than those not represented orthographically. The vowels with long duration were detected regardless of having or not having any orthographic representation. If orthographic effect was mandatory, all short vowels without orthographic representation were expected to be detected less accurately than the long vowels with orthographic representation. Therefore, orthographic knowledge could help or facilitate the task performance for the vowels with mid or short duration.

Chapter 6: Descriptive studies - Reading experiments

6-1. Introduction

By learning to read an alphabetic writing system, a link is made between constituents of written words (letters/graphemes) and the constituents of the spoken forms of words (sounds/phonemes) which leads presumably to an orthographic effect in a way that written constituents can affect the way the phonological constituents are perceived and processed (see for example, (Ehri, 2014; Rosenthal & Ehri, 2008; Ehri, 1984; Ehri & Wilce, 1980; Seidenberg & Tanenhaus, 1979)). In this respect, transparent orthographies are in contrast with opaque orthographies. In the former, grapheme-phoneme mappings are almost consistent and complete which is not always the case in the latter. This contrast provides a good opportunity to find out how the relationship between written and spoken constituents influences speech processing.

For example, English speakers report that spoken words such as ‘catch’ and ‘badge’ have more sounds than ‘much’ and ‘page’ respectively (Ehri & Wilce, 1980). Also in an auditory rhyme judgment task, rhyming words having the same rhyme spellings, such as ‘broom-room’, were judged faster than those spelled differently, such as ‘tomb-room’ (Seidenberg & Tanenhaus, 1979). Orthography can even influence the number of perceived syllables in a word. For example, speakers who knew the correct spelling of the word ‘interesting’ perceived one more syllable in it than those who spelled it as ‘*intresting’ (Ehri, 1987). There are other studies on this topic that show that processing spoken words is influenced by written words when there is a mismatch or inconsistency between sound and spelling (Cutler & Davis, 2012; Taft, 2011; Cutler, Treiman, & Van Ooijen, 2010; Dijkstra, Roelofs, & Fieuws, 1995).

So far, most studies on the effect of orthography on spoken word processing investigated effects of polygraphy (a phoneme is represented by multiple graphemes, such as /s/ represented by <c> and <s>) or of silent graphemes (graphemes that have no phonemic correspondence such as <t> in ‘catch’) to the best of the author’s knowledge, and the influence of spelling on spoken word processing has often been confirmed. In contrast, the case of the auditory processing of orthographically unrepresented phonemes has been somewhat neglected. This may be due to the nature of the languages under investigation, in which almost all sounds in the spoken words are represented in the written form of words.

Orthographically unrepresented phonemes can be found in a deep orthography such as Persian conventional writing. Persian orthography, as introduced in the first chapter, is written with a modified Arabic script and has two forms: conventional and unconventional.

The conventional writing does not cover all vowels present in the spoken language. In the conventional form, short vowels are not written, whereas long vowels are represented in writing. In the unconventional writing, however, all vowels are orthographically represented. Short vowels are represented by diacritics (◌َ, ◌ِ, ◌ُ, where circles represent the base line) above or below the base line which is used to write other letters or graphemes. The unconventional form of writing is dense, but almost shallow in this respect (مَنظَرِه, <manzareh>, /mænzæreh/, ‘scene’), whereas the conventional form is less dense but deep (منظره, <mnzrh>, /mænzæreh/) with respect to the representation of short vowels.

The unconventional form is used in the first grade of elementary school only. The conventional form is used later on for writing all types of the Persian texts. In the conventional writing, short vowel phonemes are not represented orthographically, and, unlike in Arabic, linguistic knowledge i.e., morphological knowledge about word patterns[32] can rarely help to predict short vowels in a printed word. Hence, the lexical phonological knowledge of words is the only available source for the full phonological information of printed words.

Considering the effect of orthography on spoken word processing, it may, therefore, be assumed that short vowel phonemes should be processed less accurately in orally presented words by Persian adult speakers who are monolingual and literate. The higher the proficiency at reading the conventional writing system, the poorer the short vowels should be expectedly processed in speech (for discussions related to the

---

[32] In a language such as Arabic different patterns might be applied to a root to construct different words with different meanings (Boudelaa & Marslen-Wilson, 2015). By applying a pattern, some phonemes are added to a root in certain positions. For example, د <d> /d/, ر <r> /r/, س <s> /s/ which is a root ‘drs’ can appear in different forms after the application of different patterns such as درس /**d**æ**r**æ**s**æ/ <drs> ((he) taught), دروس /**d**o**r**u**s**/ <drus> (lessons), مدرس /mo**d**æ**r**re**s**/ <mdrs> (instructor), تدریس /tæ**dr**i**s**/ <tdris> (teaching), or درس /dærs/ <drs> (lesson). It should be added that this feature does not exist in Persian.

relationship between reading proficiency and speech processing, see for example, (Saletta, 2019; Nation & Hulme, 2011; Castle, Holmes, Neath, & Kinoshita, 2003; Watson & Miller, 1993)). This effect might be due to a restructuring of phonological codes in terms of the orthographic features. Conversely, those who are better at reading the unconventional writing system may be expected to be better at processing and perception of short vowel phonemes in the auditory modality.

The reading experiments in this dissertation were mainly designed to find a possible correlation between the performance in the reading tasks and the performance in the auditory ABX discrimination task with respect to vowels. In other words, the author was interested to know whether the proficiency in reading the conventional writing correlates with the vowel discrimination ability.

Furthermore, to find how phonological features or phonological processing might be modulated in terms of orthographic features, we should have a close look at the reading processes in Persian conventional writing. It has been claimed that reading skills and cognitive processes involved in reading develop depending on how orthography encodes the spoken language (Frost, 2005; Ziegler & Goswami, 2005; Katz & Frost, 1992). Therefore, knowing how Persian readers deal with the absence of short vowels in scripts during silent reading and reading aloud might help us know how the orthographic features find their way into the phonological processing. This is the second aim of this chapter. To this purpose, the results of related studies are also considered in addition to the results of the reading experiments in this dissertation to draw conclusions.

With these aims in mind, two reading experiments were conducted. In the first one, a visual pseudoword naming task, participants were asked to name (read out aloud) written pseudowords with long vowels or pointed short vowels. In the second experiment, a semantic categorization task, participants were asked to categorize words that were written in the conventional or unconventional script as living or nonliving. In both tasks, the reading performance for short and long vowels was under focus.

There are opposing views about the role of phonology in reading. The first view (Martin & Jensen, 1988) maintains that the involvement of phonological processing in reading depends on the task requirements, and that phonological processing is not mandatory or automatic in written word recognition. Phonological codes might get activated in parallel with the visual codes for written words, but visual access occurs much faster in the recognition process. Therefore, lexical decision and visual recognition for written words are mostly performed by visual access rather than by phonological access.

Relatedly, Martin and Jensen (1988) were interested to find whether a rhyming relationship between primes and targets could facilitate the lexical decision for the written targets. They conducted a visual primed lexical decision experiment. Target written words were primed by rhyming written words with similar spelling (fool-SPOOL), rhyming written words with dissimilar spelling (rule-SPOOL), semantically related written words (thread-SPOOL), unrelated written words (waltz-SPOOL), or neutral control strings (xxxxx-SPOOL). The lexical decision latencies were the same when target words were primed by rhyming words with similar spelling, rhyming words with dissimilar spelling, or unrelated words. Lexical decision in the semantic priming condition (thread-SPOOL) was faster than the performance in the unrelated condition. Thus, the authors did not find any rhyme facilitation (phonological facilitation), whereas semantic facilitation was observed, indicating that the paradigm was sensitive enough to show priming effects. Moreover, finding a facilitatory effect for semantic priming, but not for phonological priming was taken as evidence that semantic processing is automatic in the visual lexical decision task, whereas phonological processing is not. Hence, visual lexical decision is not mediated by phonological processing.

On the other hand, others have found that phonological codes have a significant role in visual lexical decision. For example, Meyer et al. (Meyer, Schvaneveldt, & Ruddy, 1974) found that when written primes and targets shared the same sounds and spellings (e.g., bribe-tribe), lexical decision for the targets was faster than when the primes and targets were unrelated. However, when the written primes and targets shared the same spelling, but differed in their phonological form (e.g., couch-touch), no facilitation was observed. This was taken as evidence for a crucial role of phonological information in visual lexical access (see (Katz & Frost, 1992; Perfetti, Zhang, & Berent, 1992; Van Orden, 1990) for similar findings).

Accordingly, it could be assumed that printed word naming or recognition (as measured by the semantic categorization task) could be affected by spelling transparency in Persian. Assuming that phonological mediation is mandatory in written word recognition and reading, opaque written words (words with short vowels) may be expected to be read or recognized less accurately than transparent written words (words with long vowels), because the assembled phonology that is supposed to be used to access the lexicon, is not complete and does not match the addressed phonology (or the lexical phonological representations) for opaque written words in contrast to transparent words in the conventional writing.

Before explaining the details of our reading experiments, the findings of previous relevant studies are summarized in the following section. Relevant studies are those that were performed with the aim of assessing the effect of orthographic transparency on the reading performance in tasks such as written (pseudo)word naming or visual lexical decision in Arabic, Hebrew, and Persian which do not represent short vowels in the conventional writing.

6-2. Evidence from the related literature on reading

6-2-1. Arabic & Hebrew

With respect to spelling transparency defined by the grapheme-phoneme mapping, there are both commonalities and differences between Persian and Hebrew/ Arabic that need to be pointed out before reporting the literature.

First, the written symbols for short vowels /æ, e, o/ are removed from the conventional scripts of Arabic, Hebrew, and Persian, whereas long vowels are orthographically represented. Short vowels are represented (pointed) in the unconventional scripts of these languages.

Second, words in Arabic and Hebrew are derived from roots and word patterns. This is not the case in Persian. As explained earlier, the word patterns can make the presence of the short vowels predictable in the conventional writing.

Third, words receive a final short vowel depending on the syntactic roles they receive in standard Arabic. Thus, syntactic processing may help predict word-final short vowels in Arabic. This is not the case in Persian, and Persian words never end in a short vowel. In the following section, it is discussed how pointing can influence reading in Arabic and Hebrew. In other words, how pointing the script (adding short vowels to the script) can influence the accuracy and/or the speed of reading across different reading tasks.

Taha and Azaizah-Seh (2017) were interested to find out whether pointing (adding short vowel symbols) could influence written word recognition in Arabic. For that purpose, highly skilled native Arabic readers were asked to decide whether written stimuli were words or pseudowords in that language. The authors found that written word recognition was faster and more accurate for unpointed items as compared to fully or partially pointed items. They argue that the automatic reading process was interrupted by pointing the script. The authors add that in getting proficient in reading, the readers move from non-lexical (grapheme to phoneme conversion) to lexical processing. The authors interpret the results based on the local detection model of Dehaene and colleagues (Dehane, Cohen, Sigman, & Vinckier, 2005): short vowel diacritics are located outside the usual detection areas where the basic letter symbols reside. Basic symbols are those used in the conventional writing, and the perceptual mechanism of adult readers has learned to detect them to recognize written stimuli. Therefore, according to the authors, adding short vowels to the script disrupts the automatic reading process in this language. The findings of this experiment suggest that lexical access is performed by visual access without phonological mediation. If the lexical decision was mediated by phonological processing, pointing should have been an advantage because pointed strings, in contrast to partially pointed or unpointed strings, provide the full phonological information.

With respect to the effect of pointing on written word naming, Shimron (2015) reports that Hebrew readers could name pointed Hebrew words almost as accurately as un-pointed Hebrew words. However, considering that both Hebrew and Arabic provide a rich morphological structure in the form of root and word patterns, he considers pointing a redundant feature for reading in Hebrew. He mentions that pointing might facilitate naming just in certain conditions where the word pattern is uninformative. In any case, Shimron's results suggest that when there is a contribution of other linguistic knowledge, pointing does not facilitate naming written words in Hebrew.

Abu-Rabbia (1998) also studied the role of pointing in Arabic with a different experimental design. In his experiment, the texts[33] were presented in correctly pointed, unpointed, and wrongly pointed conditions. Participants (poor and skilled Arabic readers) were asked to read out the texts, and the accuracy of reading was measured across conditions. The results showed that the correctly pointed texts were read more accurately than the unpointed and wrongly pointed texts, and the unpointed texts were read more accurately than the wrongly pointed texts by both skilled and poor readers. He concluded that pointing contributed to the accuracy of the reading (naming) in both groups.

In another study searching for the effect of context and pointing in reading Arabic, the participants (poor and skilled Arabic readers) were asked to read out fully and unpointed paragraphs, sentences, and isolated words (Abu-Rabbia, 1997). There were three fixed factors: participant group (levels: skilled and poor), stimulus type (levels: word, paragraph, sentence), and pointing (levels: fully pointed and unpointed). The results of separate analyses for each stimulus type showed that paragraphs and sentences were read more accurately by skilled than by poor readers; moreover, paragraphs and sentences were read more accurately when pointed by both groups of participants. The interaction between participant group and pointing was not significant in reading paragraphs and sentences suggesting that pointing helped both groups of participants to the same extent. For words, there were also significant main effects of

[33] In Arabic, reading isolated words is different from reading the same words in a text. In texts, the words receive an extra final short vowel depending on the syntactic role they have in the text. For example, the isolated word /ræjol/ <rjl> رجل (man) can be realized with the final /o/ where it occurs as the topic or the subject (agent) of the text. The same word can be realized with the final /æ/ where it is the object (patient) in the text. Hence words are inflected for the syntactic roles they receive. All short vowels are removed in the conventional format; therefore, readers need not only identify the vowels in the isolated words; but they also need to process the syntactic structure of the text and assign the correct final vowels while reading official Arabic texts.

pointing and participant group; furthermore, the interaction between participant groups and pointing was significant indicating that the effect of pointing was more pronounced in skilled than in poor readers. In sum, the findings suggested that pointing was a good improver in reading sentences, paragraphs, or words; and that skilled readers profited more from all sources of available information while reading words.

An analysis of the data considering all three factors (participant group, stimuli type, and pointing) together showed a significant effect for the pointing only when the stimuli were words rather than paragraphs or sentences across both participant groups, suggesting that the effect of context overrode the effect of pointing in reading sentences and paragraphs. Abu-Rabbia (1997) argues that in reading texts (paragraphs/sentences), context is the only factor that contributes to correct reading. Whereas in reading words, pointing is the only factor that contributes to correct reading in Arabic.

In another study performed in Hebrew (Frost, 1995), the author investigated the role of spelling transparency in reading. In the first experiment, readers were asked to name words that contained short vowels, in which few or many short vowels were missing. Naming latencies increased with the number of the missing short vowels. In the second experiment, Frost studied the role of the number of missing short vowels in a lexical decision task. Lexical decision latencies were the same for words with few or many missing short vowels. The author suggests that the number of missing short vowels corresponds to the degrees of freedom that need to be ‘computed’ in naming but not in lexical decision tasks.

In a set of experiments in Hebrew, Shimron and Navron searched for the effect of pointing across different tasks. In a naming task (Shimron & Navon, 1982; Navon & Shimron, 1981), participants (native adults and children) were asked to name pointed and unpointed Hebrew words. All the words had one correct reading. Both children and adults named the pointed words faster than the unpointed words. The authors suggested that pointing contributed to the phonemic recoding that was required in naming. The findings also showed that experience with reading did not nullify the role of pointing as the performance of both adults and children was affected by pointing.

Later, these authors reasoned that the significant effect of pointing could be the result of the task: naming tasks may induce phonemic recoding. Therefore, they decided to test the role of pointing in tasks that did not need any pronunciation (production): Lexical decision and semantic categorization (Navon & Shimron, 1985). In the lexical decision task, they did not find any difference between the latencies for pointed and unpointed words. In the semantic categorization task, the participants were asked to judge whether the words were related to a predefined target category (e.g., if 'whale' was an animal). The results of the semantic categorization task conformed to the results of the lexical decision task: There was no significant difference between the reaction times for pointed and unpointed words. Thus, pointing contributed to the performance in the naming task only.

Koriat (1984) studied the effects of pointing and semantic priming (context) in a lexical decision task in Hebrew. He hypothesized that lexical decision could be faster for pointed strings if the process was phonologically mediated, the effect of pointing was expected to increase with the length of the stimuli. Also, the effect of pointing was expected to decrease with semantic priming (context). The results rejected the hypotheses. He found that for Hebrew words with one reading and meaning, pointing did not make the lexical decision faster. The null effect of pointing was not changed by increasing the string length; also, context did not change the null effect of pointing. In a second experiment (naming printed words), however, pointing and the length of the stimuli had significant effects on naming latencies without any significant interaction between them. Therefore, Koriat suggests that pointing affected deriving phonological codes, and stimulus length affected generating the motor commands for articulation. Thus, the results, again, suggest that naming, but not lexical decision, is influenced by pointing.

In another study, Koriat (1985) found that pointing facilitated word recognition in a lexical decision task, when low frequency words were added to the list of stimuli, suggesting that pointing can help when low frequency words are to be recognized. Koriat suggests that the strategy used for low frequency words differed from that used for high frequency words with a more strongly weighted phonological processing for the former. In a mixed list containing both low and high frequency words, participants tended to generalize the strategy suitable for the low frequency words to the high frequency words.

In consequence, the results of these experiments in Arabic and Hebrew show that access to the full phonological information of words with short vowels was demanding for Arab and Hebrew readers when the words were to be named in the conventional writing (Frost, 1995). Pointing helped access the full phonological information of written words in the naming tasks (Shimron, 2015; Abu-Rabbia, 1997; Koriat, 1984; Shimron & Navon, 1982; Navon & Shimron, 1981). Other sources of linguistic knowledge such as morphological or contextual knowledge also facilitated access to the full phonological information of opaque words in naming tasks (Shimron, 2015; Abu-Rabbia, 1997). When no production was required for task performance (e.g., in lexical decision tasks), pointing was either disrupting (Taha & Azaizah-Seh, 2017) or could not enhance the task performance (Frost, 1995; Navon & Shimron, 1985; Koriat, 1984).

6-2-2. Persian

In Persian, in contrast to Arabic and Hebrew, the presence of short vowels is not predictable as a result of the contribution of the morphological structure of written words (word patterns). Therefore, words with short vowels cannot be named by processing the script only, and the lexical phonological knowledge is the source that can lead to the correct pronunciation of printed words with short vowels in the conventional writing. This is not the case for the words with long vowels. In such words, the bottom-up processing of printed words can lead to the correct naming. The following section deals mainly with comparing the reading performance for words containing long vowels (transparent) with that for words containing short vowels (opaque) in the Persian conventional writing.

Bakhtiar and Weeks (2015) were interested to find what lexical-semantic factors could predict the performance for reading words in Persian. The lexical-semantic factors included word frequency, age of acquisition (AOA), imageability, and familiarity. They conducted a written word naming experiment with Persian transparent (words with long vowels) and opaque (words with short vowels) words in the conventional writing. Participants were asked to name the written words. The authors found that imageability, AOA, and frequency had significant main effects on naming latencies. Highly imageable, highly frequent, and early acquired words were named faster than low imageable, infrequent, and late-acquired words. There was a significant main effect of transparency indicating that transparent words were named faster than opaque words. Also, they found a significant interaction between AOA and transparency: the effect of AOA was stronger for opaque words than for transparent words suggesting that this factor had a stronger effect on naming opaque words than on naming transparent words. Moreover, the three-way interaction between AOA, imageability, and transparency was significant indicating that the effect of AOA was larger for opaque word with low imageability than for opaque words with high imageability. Accordingly, the authors suggest that reading late acquired opaque words received more input from the lexical semantic route. Considering the DRC Model for written word processing (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), the authors conclude that opaque words were named by the lexical-semantic route, whereas naming transparent words was performed by the non-lexical route.

In another naming study, Rahbari & Senechal (2008) found that in the conventional writing, transparent words were named faster than opaque words. Naming opaque and transparent words was faster than naming non-words. Naming in both conditions (‘opaque’ and ‘transparent’) was affected by word frequency. Since word frequency is a lexical property, they came to the conclusion that Persian readers follow the lexical route in the DRC Model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) for naming both opaque and transparent words. As for the accuracy of naming in Persian, Rahbari (2019) found that naming opaque written words was less accurate than naming transparent written words in the Persian conventional writing. Thus, opaque words are named more slowly and less accurately than transparent words.

In a series of naming experiments with Persian native speakers, Baluch and Besener (1991) found that word frequency and semantic priming influenced naming speed for opaque as well as transparent words when the stimuli were real words only. However, word frequency and semantic priming influenced the naming speed for opaque words only, when nonwords were included in the stimulus list. The significant effect of the lexical-semantic factors (word frequency and semantic priming) suggests that both transparent and opaque words were named by lexical or lexical-semantic routes in the list without nonwords (considering the DRC Model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001)). But by including nonwords in the experimental context, different strategies were used to name opaque and transparent words. Transparent words were named by the non-lexical route, whereas the opaque words were named by the lexical or lexical-semantic routes. In their conclusion, the authors emphasize the flexible use of lexical and non-lexical strategies in the speeded naming of transparent words depending on the experimental context.

Comparing the results of the experiments by Bakhtiar and Weeks (2015) with those by Baluch and Besener (1991), Rahbari (2019), and Rahbari & Senechal (2008), it becomes clear that Persian transparent written words can be named by either the lexical or the non-lexical route, or even by both routes simultaneously. The main strategy is dependent on the types of stimuli and the design of the experimental task. In contrast, opaque written words are necessarily named by the lexical route. Thus, transparent written words have more routes available than opaque written words in naming tasks; and this point can be considered as an advantage for transparent written words (Bakhtiar & Weekes, 2015).

The advantage of spelling transparency is not limited to written word naming studies cited above. Even in a recall study (Baluch & Dana-ye-Tousie, 2006), the trend was the same. Baluch and colleague designed a two-part experiment. In the first part, the participants were asked to name words manipulated for frequency (high vs. low), imageability (high vs. low), and spelling (transparent vs. opaque) presented in the conventional writing. After the naming task, the participants were asked to write as many words as they could recall from the preceding experiment (naming). The factors (transparency, frequency, and imageability) were all ineffective in the naming performance, but had significant effects in the recall test with significant two- and three-way interactions indicating that transparent words were recalled better when they were highly frequent-highly imageable or infrequent-low imageable compared to opaque words with the same properties. Baluch and Dana write that the results were against their expectation. They had expected opaque words to be recalled better and faster due to the high lexical involvement during naming, which might have left a stronger memory trace. Instead, they found the reverse i.e., transparent words were recalled better than opaque words suggesting an advantage of spelling transparency.

Considering the experiments without naming component, in a visual lexical decision task, Baluch (1993) found no significant difference between the reaction times for high frequency transparent and opaque words. For low frequency words, he found that transparent words were responded to faster than opaque words, suggesting that in lexical decision, different strategies may be used for high and low frequency words. Note that spelling transparency is an effective factor in the lexical decision with low frequency written words and in naming written words (the experiments cited above), it is likely that the same strategy has been adopted in the lexical decision for low frequency words and in naming words.

To the best of the author's knowledge, there is only one published experiment on the effect of pointing the short vowels on reading in Persian. Hosseini Almadani (2016) measured the proficiency of bilingual (Persian (L1)-English) children in a naming task. The participants (7-14 years) were asked to name pointed and unpointed words with short vowels. She found that Persian bilinguals were better at naming pointed than unpointed words. The author concludes that pointing can facilitate naming. Even having more experience in reading did not change the effect of pointing. The result of this experiment is in line with the results of the naming experiments in Arabic and Hebrew (see above).

In sum, considering the literature on Persian, spelling transparency is an influential factor in naming experiments and other tasks inducing phonological recoding (e.g., lexical decision with low frequency words), whereas lexical decision on written words seems to be independent of spelling transparency.

All in all, the literature on Persian and other languages (Arabic and Hebrew) showed that spelling transparency in the case of vowels had an influence on the performance in tasks with phonological output, but not on the performance in tasks without phonological output. Moreover, pointing short vowels improved the performance in tasks with phonological output, but not in tasks without phonological output.

6-3. Experiments

6-3-1. Naming experiment

6-3-1-1. Participants

The literate participants in the auditory ABX discrimination task were asked to participate in this experiment.

6-3-1-2. Stimuli

Forty-eight pseudowords (P) were constructed in a way to contain just one long vowel (PL1, e.g., <pal> /pɑl/), three long vowels (PL3, e.g., <kinubas> /kinubɑs/), one short vowel (PS1, e.g., <mol> /mol/), or three short vowels (PS3, e.g., <talesom> /tælesom/). Those that contained just one vowel had a CVC or CVCC structure; and the pseudowords with three vowels had a CVCVCVC structure. The vowels in the long pseudowords were different. Twelve pseudowords were included in each condition and all were consistent with the phonotactics of Persian.

The pseudowords were constructed by changing at least one consonant of Persian real words. Note that changing vowels to create pseudowords would have affected the written representations of words with long and short vowels differentially. If, for example in a real word with a short vowel such as /særd/ <srd> سرد in the conventional writing, the short vowel was changed to create a pseudoword from that word, the overall look for the real word and the created pseudoword would have stayed almost the same, e.g., the pseudoword: /serd/ <serd> سِرد. By contrast, this would not have been the case for written words with long vowels. If, for example, in a real word with a long vowel such as /bɑr/ <bar> بار, the long vowel was changed to create a pseudoword, the resulting string would be /bir/ <bir> بیر with a different overall look of the real word and the pseudoword. Therefore, to have an equally modified visual form for pseudowords with short or long vowels, consonant(s) rather than vowels of the real words were changed to create the pseudowords.

The written stimuli with short vowels were presented in the pointed format. All stimuli were pseudorandomized for order and condition. Stimuli were arranged in three different orders across three different lists, and stimuli from the same condition did not immediately follow each other within each list. At the beginning of a list, 10 warm-up items were included.

Ideally, the reading experiments should have included naming both real and pseudowords. However, as the experimental sessions per participant already took more than four hours (ABX, phoneme monitoring, naming pseudowords, semantic categorization for real words, and trainings), the experimenter decided to shorten the naming experiment by only including pseudowords and to rely on the findings of other studies on naming Persian real words for comparisons between naming pseudowords and real words.

6-3-1-3. Procedure

Each participant was tested individually. The experimenter explained the procedure orally before the commencement of the experiment. When participants were ready, the experiment started on a laptop screen. At the beginning, a written instruction, asking participants to name the visual pseudowords as fast as and as accurately as possible, appeared on the screen and remained there to be fully read and understood. As soon as the participants pressed the space button after reading the instruction, strings of pseudowords were displayed one by one on the screen to be named.

Each trial started with a plus sign displayed at the center of the screen for 200 *ms* to show the place where the visual pseudoword was going to be displayed. 100 *ms* later, the pseudoword was displayed for 1000 *ms* on the screen, and the participants had 2500 *ms* time from the pseudoword onset to name it. The microphone (voice key) was open during this time to record the participants' outputs. The inter trial interval was set to 1000 *ms*. At last, a 'Thank you' note indicated the end of the naming test.

A personal Sony laptop with a 17" monitor was used to present the written stimuli for the experiment. Responses were recorded as audio files on the laptop for each participant. Praat 5.4.09 (Boersma & Weenink, 2015), was used to analyze the accuracy of the output data.

The experiment was programmed and presented in Psychopy V.1.82.01 (Peirce, Psychopy, 2002-2018). Because this software does not support Persian fonts, the pseudowords were typed in a PowerPoint file (font 'Nazanin', size 24"), and then converted to image files (image size '960*720') to be presentable in Psychopy. All stimuli could be easily read from a 50-70 *cm* distance from the monitor display, as judged by three Persian native speakers. All participants reported unimpaired or appropriately corrected eyesight, no motor problem, and no problem in speech production.

6-3-1-4. Data analysis

The output data were analyzed for the accuracy of the vowel readings. Problems with consonants reading were ignored, and were not counted as errors. Incorrect replies were coded in two different ways:

1) the accuracy was measured per pseudoword (left panel, Figure 17). A misreading of the pseudoword vowel/ vowels was counted as one incorrect reply, irrespective of the number of misread vowels within a pseudoword.

2) the accuracy was measured per vowel position within pseudowords (right panel, Figure 17). In this approach, the accuracy of each position was scored separately within pseudowords to assess the vulnerability of each position to misnaming. Vowel position referred to the respective position within the pseudowords where a vowel occurred. For example, for the pseudowords in the PL3/PS3 conditions, vowels occurred in three positions $CV^1CV^2CV^3C$. Here the superscripts stand for the respective vowel position. The error rates for each position within a pseudoword were calculated independently of other positions in this approach. If a participant misnamed all the vowels in a pseudoword with three vowels, this was scored as three errors in naming that pseudoword, one per each position.

Items were coded as 'not-replied', when the participants could not read them at all or read them partially (9.4% of the whole dataset). Not-replied items were removed from further analysis, since failing to name the pseudoword could not be attributed solely to a problem with vowel reading. The number of items across corresponding conditions remained almost matched after removing the not-replied items.

General linear mixed effect regression analyses (lem4 package in R (Bates, Mächler, Bolker, & Walker, 2015)) were used to analyze the accuracy of naming. Two models were used. In the first model, 'pseudoword type' (PL1, PS1, PL3, PS3) was considered as the fixed factor. In the second model, the fixed factors were 'pseudoword type' and 'vowel position'. In both models, the random factors were 'participants' and 'items'. The effect of the fixed factor was assessed by comparing the model including all factors with the model without that target factor (ANOVA tests). The multcomp package (Tukey test (Hothorn, Bretz, & Westfall, 2008)) was used for pairwise comparisons between the levels of the fixed factors. The reported *p* values were corrected for the number of comparisons.

6-3-1-5. Results

A detailed analysis of the replies showed that in incorrect replies existing vowels were replaced by the vowels of the same category. That is, the existing short vowel was replaced by another short vowel, but never by a long vowel. Short vowels were never omitted to produce strings that were phonotactically illegal. Figure 17 (left panel) shows the mean proportion of incorrect replies (error) as a function of 'pseudoword type'.

The mean proportion of errors was M=0.02, SD=0.05 for naming written pseudowords with one long vowel and M=0.07, SD=0.11 for naming written pseudowords with one short vowel. The mean proportion of errors was M=0.23, SD=0.2 for naming pseudowords with three long vowels and M=0.66, SD=0.2 for naming written pseudowords with three short vowels. Hence, participants made about three times more errors in naming written pseudowords with short vowels as compared to naming written pseudowords with long vowels[34].

The results of the general linear mixed effect regression to analyze naming accuracy (correct vs. incorrect) as a function of 'pseudoword type' (PL1, PL3, PS1, and PS3) showed that 'pseudoword type' had a significant effect on the accuracy of naming in this experiment [$X^2(3)$ =94.9, p<0.0001]. Pairwise comparisons (using Tukey tests)

---

[34] The mean reaction times for correct items measured from the onset of the pseudoword display to the onset of the naming were PS1=1541 *ms*, PL1=1495 *ms*, PS3=1705 *ms*, PL3=1567 *ms*.

between the corresponding levels of 'pseudoword type' showed that pseudowords with one long vowel were named as accurately as pseudowords with one short vowel (Z=2.2, p=0.1). However, pseudowords with three short vowels were named less accurately than pseudowords with three long vowels (Z=7.6, p<0.001).

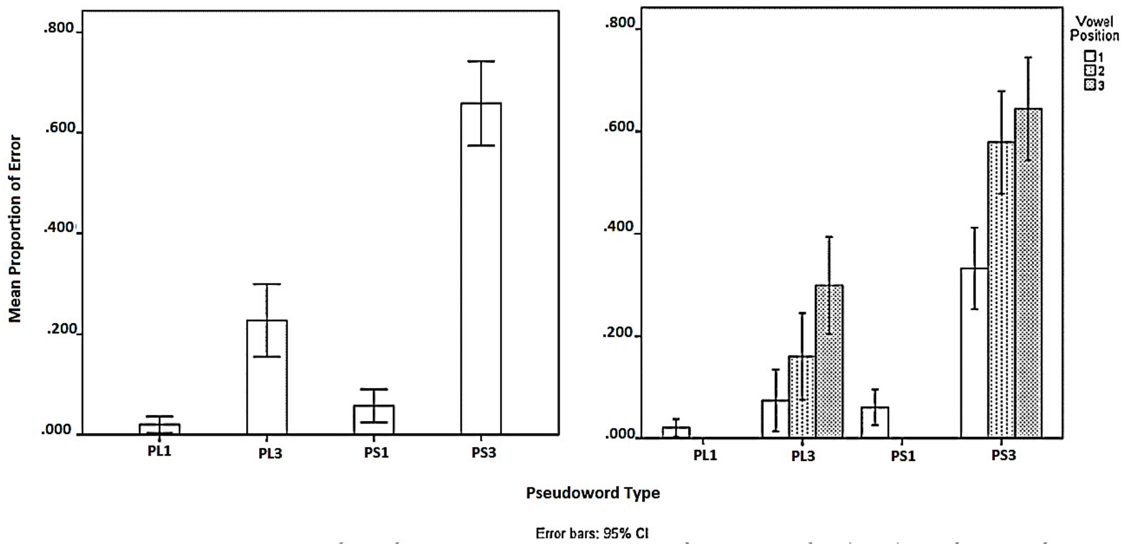


Figure 17. Naming written pseudowords in Persian. Mean proportion of incorrect replies (error) as a function of 'pseudoword type' (left panel). Mean proportion of incorrect replies (error) as a function of 'pseudoword type' and 'vowel position' (right panel).

The analysis of naming accuracy as a function of 'pseudoword type' and 'vowel position' (Figure 17, right panel) showed that pseudowords with three short vowels were named with higher mean proportions of errors as compared to pseudowords with three long vowels across all vowel positions (1, 2, and 3). In the first vowel position, the mean proportion of errors was M=0.07, SD=0.17 for pseudowords with three long vowels ($PL3^1$) and M=0.33, SD=0.23 for pseudowords with three short vowels ($PS3^1$). In the second vowel position, the mean proportion of errors increased from M=0.16, SD=0.28 for pseudowords with three long vowels ($PL3^2$) to M= 0.58, SD=0.25 for pseudowords with three short vowels ($PS3^2$). In the third vowel position, the mean proportions of errors were M=0.3, SD= 0.29 and M=0.64, SD=0.25 respectively for pseudowords with three long vowels ($PL3^3$) and pseudowords with three short vowels ($PS3^3$).

The results of the general linear mixed effect regression for the analysis of the accuracy of naming as a function of 'pseudoword type' and 'vowel position' (1, 2, 3) showed significant main effects of ‘pseudoword type’ [$X^2(3) = 301$, $p<0.001$] and ‘vowel position’ [$X^2(4) = 143$, $p<0.0001$] on the accuracy of naming. The interaction between ‘pseudoword type’ and ‘vowel position’ was also significant [$X^2(2) = 6$, $p=0.03$].

Pairwise comparisons between the corresponding levels of ‘vowel position’ across ‘pseudoword type’ ($PL3^1$-$PS3^1$; $PL3^2$-$PS3^2$; $PL3^3$-$PS3^3$) showed significant differences across all levels (position 1: $Z=-7.43$, $p<0.0001$; position 2: $Z=-10.74$, $p<0.0001$; position 3: $Z=-8.95$, $p<0.0001$). Comparing the accuracy of naming for the first vowel, the result showed that the first short vowel was named less accurately in the long pseudowords ($PS3^1$) than in the short pseudowords ($PS1^1$) ($Z= -11$, $p<0.001$). But the first long vowel was named with the same accuracy in the long pseudowords ($PL3^1$) and in the short pseudowords ($PL1^1$) ($Z=-2.13$, $p=0.9$). Also comparing the accuracy of naming across different positions within vowel category, the results showed that the first long vowel was named as accurately as the second long vowel ($Z=-2$, $p= 0.2$), but better than the third long vowel ($Z=-6.8$, $p=0.03$) in long pseudowords. On the other hand, the first short vowel was named more accurately than the second ($Z=-6.5$, $p=0.01$) and third ($Z=-8.37$, $p=0.01$) short vowels, while the naming accuracy remained the same between the second and third short vowels ($Z=-2.13$, $p=0.2$) in long pseudowords.

6-3-1-6. Discussion

The results of the naming experiment showed that naming written pseudowords with one pointed short vowel was as accurate as naming written pseudowords with one long vowel. However, for longer pseudowords (including three vowels), pseudowords with pointed short vowels were named less accurately than pseudowords with long vowels. Furthermore, the analysis of the accuracy of naming as a function of ‘pseudoword type’ and ‘vowel position’ showed that the error rates were significantly higher for pseudowords with three pointed short vowels than for pseudowords with three long vowels in all corresponding positions.

The findings showed that pseudoword naming accuracies were affected by the category of vowels included in pseudowords, the length of pseudowords, and the position of critical vowels. Vowels in later positions were named less accurately, while misnaming in later positions was more pronounced in pseudowords with three pointed short vowels than in pseudowords with three long vowels. This result suggests that the memory span for pseudowords containing short vowels was lower than the memory span for pseudowords containing long vowels. This finding is in line with the findings of Baluch and Dana-ye-Tousie (2006) for Persian real words with long vs. short vowels. In their study, the recall rate was lower for words with short vowels than for words with long vowels after a written word naming task.

Further evidence that the memory span for long vowels was better than that for short vowels is provided by the observation that the first vowel in pseudowords with three short vowels ($PS3^1$) was named less accurately than the vowel in pseudowords with one short vowel ($PS1^1$). By contrast, the first vowel in pseudowords with three long vowels ($PL3^1$) was named as accurately as the vowel in pseudowords with one long vowel ($PL1^1$)[35]. This result suggests that the participants started naming only after analyzing the whole letter string of pseudowords and the corresponding sounds; and by the time of production (naming), the memory for the first vowel had decayed in pseudowords with three short vowels (PS3), but not in pseudowords with three long vowels (PL3).

Moreover, the comparison between the accuracy of naming across vowel positions within vowel category ($(PL3^1= PL3^2) \neq PL3^3$ and $PS3^1 \neq (PS3^2 = PS3^3)$) showed that participants had problems naming the third short or long vowel. Also, the first two long vowels were named more accurately than the last long vowel, whereas the first short vowel was named more accurately than the last two short vowels in long pseudowords. That means more vowel-positions are susceptible to misname in pseudowords with pointed short vowels compared to pseudowords with long vowels suggesting that the recall rate was better for long vowels as compared to short vowels.

The findings of the pseudoword naming experiment showed that Persian readers had no problem identifying short vowel diacritics, because naming pseudowords with

[35] It should be noted that by the time the participants started naming, the pseudowords were not on the screen.

one short vowel was as accurate as naming pseudowords with one long vowel. But the long vowels could be processed and named better than the short vowels in the long pseudowords suggesting that when the script was more familiar in pseudowords with long vowels, it could be processed more easily leading to saving time and resource to keep the constituents better in memory in pseudowords with long vowels. However, when the script was less familiar in pseudowords with pointed short vowels, the script processing could be more demanding leading to less time and resource to keep the constituents in memory in pseudowords with pointed short vowels.

Higher error rates for the later positions especially in pseudowords with pointed short vowels as compared to pseudowords with long vowels can also be explained by an interference effect known from working memory research (Baddeley, 2012): the higher the number of intervening items, the lower the accuracy of naming or recalling items in later positions. This effect was more salient for pseudowords with pointed short vowels compared to pseudowords with long vowels.

Alternatively, it might be assumed that Persian readers might be phonemically less aware of short vowels compared to long vowels. Phonemic awareness refers to the ability to attend to sound structures and to perform mental operations on them. Phonemic awareness has been shown to be strongly related to the accuracy of naming pseudowords (González-Valenzuela, Díaz-Giráldez, & López-Montiel, 2016). After processing visual symbols, the process dedicated to phonological recoding and assembling sounds to utter pseudowords is assumed to be affected by phonemic awareness (González-Valenzuela, Díaz-Giráldez, & López-Montiel, 2016). According to these authors, phonological processing is performed more accurately in transparent than in opaque orthographies due to higher phonemic awareness in transparent orthographies. Therefore, lower awareness to short vowels, compared to long vowels, could explain higher error rates in naming pseudowords with pointed short vowels as compared to pseudowords with long vowels. The results of the phoneme reversal and phoneme monitoring tasks in Chapters three and five suggest lower awareness of Persian readers to short vowels as compared to long vowels.

The findings of the pseudoword naming task showed that Persian readers had problems naming pseudowords with more than one short vowel, although the short vowels were pointed; and, transparency was not an issue. This could be because sounds with orthographic representations can be processed and retained in memory more efficiently than sounds without orthographic representations after a long exposure to the Persian conventional writing, or phonological awareness to sounds represented orthographically is higher than phonological awareness to sounds that are not spelled in the conventional writing. In other words, the memory span or phonological awareness to short vowels may have been reduced in Persian readers as a result of the long exposure to the Persian conventional writing.

6-3-2. Semantic categorization experiment

In this task, it was of interest to know whether spelling transparency had any influence on the semantic categorization of Persian written real words. Spelling transparency was defined as the contrast between short vowels without orthographic representations and long vowels with orthographic representations.

In this task, the participants were asked to categorize whether written words named living beings or nonliving things, and press the response keys accordingly. The experiment was performed in two separate blocks: The first block contained words in the conventional writing, where only long vowels were orthographically represented. In the second block, the written words with short vowels were presented in the pointed format (e.g., بَند, <band>, /bænd/ ‘string’). In this unconventional writing used for first graders only, short vowels have their own orthographic representations that are left out from the second class onwards. Hence, a written word that is initially represented as بَند, <band> /bænd/ becomes بند, <bnd> /bænd/ from the second grade onwards.

Assuming that devoting less resources to the lexical processing of words could save resources or time for the higher-level processing required for semantic categorization, we hypothesized that, due to their transparent spelling, written words with long vowels might be categorized with a higher accuracy than written words with short vowels in the conventional writing. Likewise, pointed written words with short vowels might be categorized more accurately than unpointed written words with short vowels. The written words were manipulated for word length (one vowel, three vowels) and word frequency (high frequency words, low frequency words) in addition to spelling transparency to gain more insight on the possible procedure while performing the task. Any difference between the performance for the short and long written words could, for example, show some sort of the covert subvocalization (Koriat, 1984) or computation of orthographically unrepresented vowels (Frost, 1995) while performing the task (see Section 6-2-1 for the effect of word length on generating motor command or on computation of orthographically unrepresented short vowels while reading written words). The written words were also manipulated for frequency. It has been shown that lexical identification process which precedes meaning retrieval are frequency sensitive, so that lexical access is more accurate/faster for high frequency words than for low frequency words in visual word recognition tasks (Monsell, Doyle, & Haggard, 1989). Consequently, the semantic categorization was generally expected to be more accurate for high frequency written words as compared to the low frequency written words. Moreover, it could be the case that the spelling transparency would insert its effect differentially in words with different lengths or with different frequencies, thus, word length and word frequency were controlled.

6-3-2-1. Participants

The same literate participants, who took part in the ABX discrimination, phoneme monitoring, and naming tasks, were asked to participate in this experiment.

6-3-2-2. Stimuli

Ninety-seven words were selected for the first block of the semantic categorization task. These words were manipulated with respect to their frequency (high frequency (HF) vs. low frequency (LF)), category of their vowels (short vowel (S) vs. long vowel (L)), and their length (short word vs. long word). The words in this block were presented in the conventional writing. The combination of these properties yielded eight conditions which are shown in Figure 18 and Table 5.

Additionally, forty-eight words containing short vowels were selected for the second block of the semantic categorization task. These words were manipulated with respect to their frequency (high frequency (HF) vs. low frequency (LF)), and their length (short word vs. long word). The combination of these properties yielded four different conditions which are shown in Figure 18 and Table 5. The words in this block were presented in the pointed format.

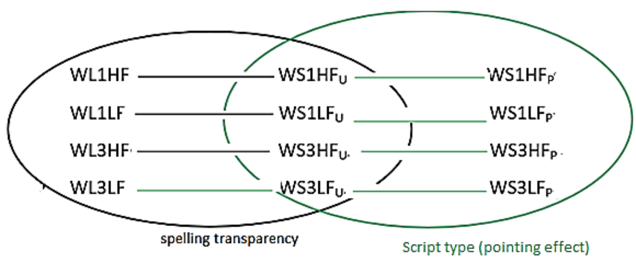


Figure 18. Conditions, grouping, and corresponding levels in the semantic categorization experiment with written words. Abbreaviations: W=word, L=long vowel(s), S=short vowel(s), 1=one vowel, 3=three vowels, HF=high frequency, LF=low frequency, u=unpointed, p=pointed.

Some of the words in each condition named living beings and the rest named nonliving things. The participants were asked to decide whether a written word referred to a living being or a nonliving thing and press the response key accordingly. The number of the words referring to living beings was not equal to the number of the words referring to nonliving things due to a shortage in finding suitable words which fit the condition requirements for frequency, category of the vowel, and number of the vowels within words. Therefore, seventeen words were additionally included as fillers to make the number of living beings equal to the number of nonliving things. The fillers were not included in the statistical analysis.

The frequency of the words was extracted from the Google engine searching in Persian pages (July 2017). For example, the word کتاب, <ktab>, /ketɑb/ 'book' was found 25,600,000 times in the Persian pages. Words with a frequency higher than twenty million were considered as high frequency words. Words that occurred less than eleven million times were considered as low frequency words. Since the words are not pointed in the conventional writing, the frequencies of un-pointed words were considered for the pointed words in the second block. The reported frequencies in the Google search were divided by the frequency reported for the most frequent word in Persian /dær/ 'in, door' (relative frequencies). The mean relative frequencies for the words in each condition is reported in Table 5.

All the experimental items were unambiguous, having exactly one pronunciation and one meaning. All the stimuli in each block were pseudorandomized; so that the number of the same consecutive correct response was limited to two, and the words from the same condition did not immediately follow each other within the list. Also, three different lists were prepared across which the order of the items differed, and each list was assigned to 1/3 of the participants. Ten warm-ups were put at the beginning of the list for each block to make the participants familiar with the procedure. The performance on warm-up trials was excluded from the statistical analysis. The words were presented in the same size, font, and programming software that were used for the pseudoword naming task.

Table 5. Characteristics of the stimuli in the semantic categorization experiment with written words. Abbreviations: No.: number, Freq.: frequency, Avg.: Average.

Block	Conditions	Abbreviation	No. stimuli	Living/ nonliving	Example			Avg. relative Freq.
					spelling	sound	meaning	
First	High frequency words with one long vowel	WL1HF	13	5/8	<dig>	/dig/	Pot	0.1
	Low frequency words with one long vowel	WL1LF	14	5/9	<ghush>	/ɣuʃ/	Falcon	0.009
	High frequency words with three long vowels	WL3HF	13	7/6	<vitamin>	/vitɑmin/	Vitamin	0.02
	Low frequency words with three long vowels	WL3LF	12	4/8	<?abdarchi>	/?ɑbdɑrʧi/	Teaboy	0.0045
	High frequency words with one short vowel (un-pointed)	WS1HF(u)	10	6/4	<zn>	/zæn/	Woman	0.12
	Low frequency words with one short vowel (un-pointed)	WS1LF(u)	12	5/7	<shykh>	/ʃeyX/	doyen	0.007
	High frequency words with three short vowels (un-pointed)	WS3HF(u)	9	4/5	<ghfsh>	/ɣæfæseh/	shelf	0.02
	Low frequency words with three short vowels (un-pointed)	WS3LF(u)	14	8/6	<vrghh>	/væræɣeh/	sheet	0.004
Second	High frequency words with one short vowel (pointed)	WS1HF(p)	12	6/6	<kafsh>	/kæfʃ/	shoe	0.09
	Low frequency words with one short vowel (pointed)	WS1LF(p)	12	5/7	<mosht>	/moʃt/	fist	0.008
	High frequency words with three short vowels (pointed)	WS3HF(p)	10	3/7	<barzegar>	/bærzegær/	farmer	0.027
	Low frequency words with three short vowels (pointed)	WS3LF(p)	14	6/8	<sepahbod>	/sepæhbod/	Lieutenant	0.003

6-3-2-3. Procedure

As explained earlier, this task had two blocks. In the first block, participants were asked to categorize written words with short or long vowel(s) in the conventional writing as living beings or nonliving things. After a short break, the words with short vowel(s) in the unconventional writing (pointed) were semantically categorized.

Each participant was tested individually in a quiet room. First, the experimenter explained the procedure orally by giving some examples, and familiarized the participants with pressing the response key while looking at the screen. Afterwards, the participants were asked to follow the materials on the laptop screen.

First, a written instruction asking participants to press the respective response key as fast as and as accurately as possible, appeared on the screen and remained there to be fully read and understood. When ready, participants pressed the space button for the experiment to begin. A trial began with a plus sign shown for 200 *ms* at the center of the screen where the word was to be displayed. 100 *ms* after the plus sign, a word appeared on the screen and remained there for 1000 *ms*. From the onset of the written word display, the participants had 2000 *ms* time to respond via the keyboard buttons whether the word was a living being or a nonliving thing. The inter stimuli interval was set to 1000 *ms*. At the end, a 'Thank you' note indicated the end of the experiment.

The participants were instructed to press the left arrow key on the laptop keyboard for the living beings, and press the right arrow key for the nonliving things. The key assignment was reversed for half of the participants (left key for nonliving things and right key for living beings). The replies were recorded in an excel output file for each participant.

6-3-2-4. Data analysis

The output data were analyzed by SPSS (IBM Corp., 2012) and R (R Core Team, 2014). Missing responses (10% of the trials) were excluded from the statistical analysis. The number of the stimuli in each condition remained matched after deleting the not-replied items (missing responses). The mean proportions of incorrect replies (errors) are reported in the graphs (Figure 19).

The graphs were generated by SPSS (IBM Corp., 2012), and the inferential statistics were performed using R (R Core Team, 2014). General linear mixed effect regression analyses (lme4 package (Bates, Mächler, Bolker, & Walker, 2015)) were used to analyze the accuracy of the data. The main effect of each factor was calculated by comparing the model considering all factors with the model without that target factor using ANOA tests. Pairwise comparisons were performed using the multcomp package (Tukey test (Hothorn, Bretz, & Westfall, 2008)) with the reported *p* values adjusted for the number of comparisons. The random factors 'participants' and 'items' were inserted into all models.

In the semantic categorization task, the performances were analyzed to find whether spelling transparency and pointing influenced the performance. To find the effect of spelling transparency, the performance for written words with long vowels was compared against the performance for written words with short vowels in the conventional writing. In this analysis, the fixed factors were 'vowel category' (long vs. short), 'word length' (short word vs. long word), and 'word frequency' (high frequency vs. low frequency). The 'vowel category' represented spelling transparency.

To find the effect of pointing, the performance for pointed written words with short vowels was compared against the performance for un-pointed written words with short vowels. In this analysis, the fixed factors were 'script type' (pointed (p) vs. un-pointed (u)), 'word length' (short word vs. long word), and 'word frequency' (high frequency vs. low frequency). In Figure 18, the left ellipse and lines show the model for the effect of the spelling transparency; and the right ellipse and lines show the model for the effect of pointing. Corresponding levels are linked by the lines in each model.

6-3-2-5. Results

Figure 19 (top) shows the mean proportion of incorrect replies (errors) as a function of 'vowel category', 'word length', and 'word frequency'. The results of the semantic categorization in the conventional writing showed that the mean proportion of errors was lower in high frequency words with long (M_{WL1HF}= 0.18, SD=0.03; M_{WL3HF}=0.23, SD=0.025) or short (M_{WS1HF}= 0.16, SD=0.04; M_{WS3HF}=0.2, SD=0.03) vowels than in low frequency words with long (M_{WL1LF}= 0.31, *SD*=0.02; M_{WL3LF}=0.3, *SD*=0.033) or short (M_{WS1LF}=0.26, *SD*=0.02; M_{WS3LF}=0.33, *SD*=0.033) vowels.

The results of the general linear mixed effect regression analysis for the conventional writing showed no significant effects of word length [$X^2(4)$ =2.42, p=0.66] and vowel category [$X^2(4)$ =2.05, p=0.73], but a significant main effect of word frequency [$X^2(4)$ =11.7, p=0.02] on the accuracy of the semantic categorization. There were no significant two-way [vowel category * word frequency: $X^2(1)$ =0.17, p=0.68; vowel category * word length: $X^2(1)$ =0.49, p=0.49; word frequency * word length: $X^2(1)$ =0.35, p=0.56] or three-way [$X^2(4)$ = 1.55, p=0.82] interactions between the factors[36].

Figure 19 (bottom) shows the mean proportion of errors as a function of 'script type' (pointed (p) vs. un-pointed (u)), 'word length', and 'word frequency' for words with short vowel(s). The mean proportions of errors for high frequency words with pointed short vowel(s) are $WS1HF_P$: *M*=0.26, *SD*=0.032 and $WS3HF_P$: *M*=0.2, *SD*=0.03 and for low frequency words with pointed short vowel(s) are $WS1LF_P$: *M*=0.2, *SD*=0.03 and $WS3LF_P$: *M*=0.33; *SD* = 0.036.

The results of the general linear mixed effect regression analysis showed no significant main effects for 'script type' [$X^2(4)$ =3, p=0.56] and 'word length' [$X^2(4)$=5.6, p=0.23], and a marginally insignificant effect for 'word frequency' [$X^2(4)$=8.7, p=0.07] on the accuracy of semantic categorization. There were no significant two-way [script type * word length: $X^2(1)$ =0.05, p=0.82; script type * word

[36] As for the reaction time analyses: the reaction times were measured from the onset of the word display to the response time. The reaction times to correct replies across conditions in the conventional writing: WS1HF: 1504 *ms;* WS1LF: 1513 *ms;* WS3HF: 1477 *ms;* WS3LF: 1533 *ms;* WL1HF: 1494 *ms;* WL1LF: 1490 *ms;* WL3HF: 1515 *ms;* WL3LF: 1531 *ms.*
From the fixed factors, the main effect of word length was significant [$X^2(4)$ = 11.15, p=0.025], and other factors did not have any significant effect on the reaction times to correct replies in the semantic categorization of written words in the conventional writing (all *ps*>0.2). This significant main effect was driven by low frequency words with long vowel(s) (t=2.83, p=0.005).

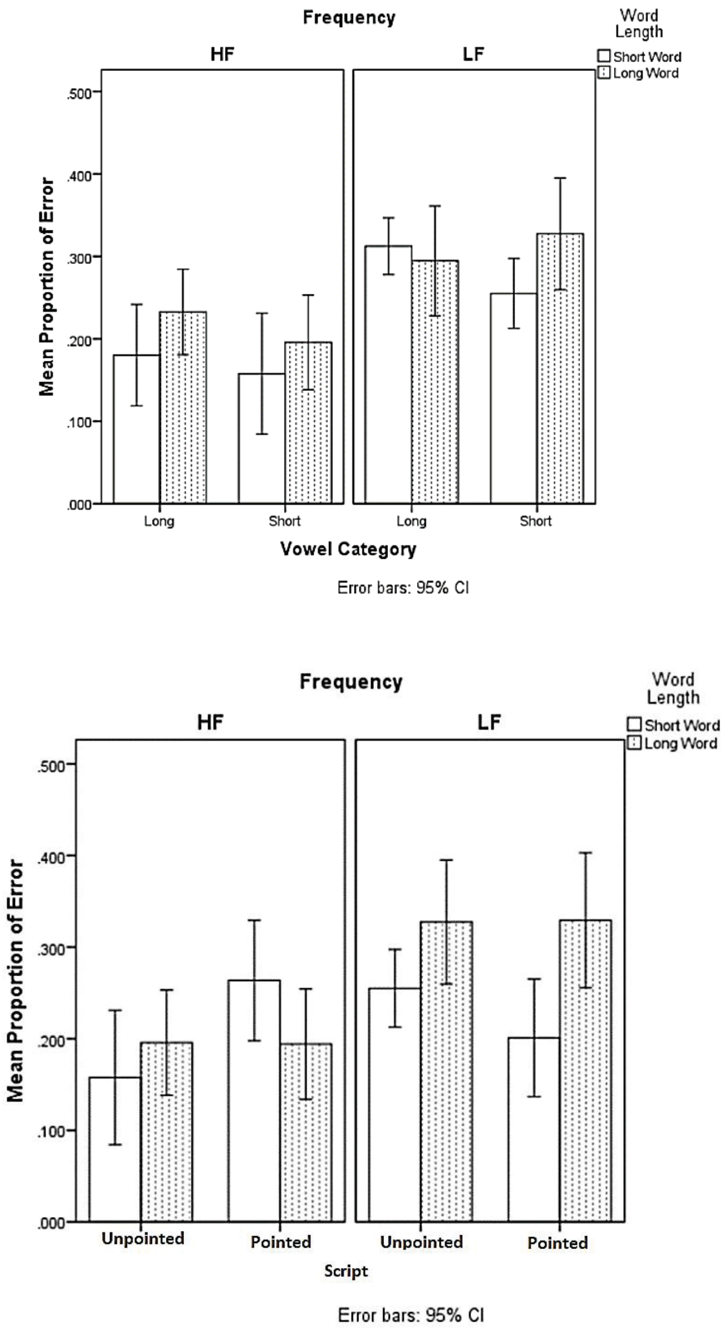


Figure 19. Semantic categorization of Persian written words. Mean proportion of incorrect replies (error) as a function of vowel category and word length for high frequency (HF) and low frequency (LF) words (top panel). Mean proportion of incorrect replies (error) as a function of script (pointed, unpointed) and word length for high frequency (HF) and low frequency (LF) words with short vowels (bottom panel).

frequency: $X^2(1) = 1.63$, p=0.2; word frequency * word length: $X^2(1) = 2.8$, $p=0.09$] or three-way [$X^2(4) = 5.7$, $p=0.22$] interactions between or among factors[37].

6-3-2-6. Discussion

Considering that short vowels are not represented in the conventional writing of Persian, whereas long vowels are, we hypothesized that semantic categorization for transparent written words would be more accurate than that for opaque written words, because more resources might be needed for the lexical identification of opaque words, and less resources might be left for higher level (i.e., semantic) processing. Conversely, the transparency of spelling for words containing long vowels might save more resources for semantic categorization. For the same reason pointing might improve the semantic categorization of opaque words.

Whereas the results of a recall experiment by Baluch and Dana (Baluch & Dana-ye-Tousie, 2006) were in line with these predictions, i.e., the recall rate was lower for written words with short vowels as compared to that for the written words with long vowels, the results of our semantic categorization experiment did not confirm them. The accuracy of the semantic categorization was not dependent on the words' spelling transparency in the conventional writing. Furthermore, inserting short vowel diacritics in the script to make the words with short vowels orthographically transparent did not lead to a better performance: the accuracy was the same for pointed and un-pointed words with short vowels in the semantic categorization experiment. The null effect of spelling transparency did not change across written words with different frequencies or lengths.

The findings of our semantic categorization task are, however, in line with a number of previous studies (e.g., (Baluch, 1993; Navon & Shimron, 1985; Koriat, 1984;

---

[37] The reaction times for correct replies in the pointed scripts are as follows. They were measured from the onset of the word display to the response time:
$WS1HF_p$: 1598 *ms;* $WS1LF_p$: 1587 *ms;* $WS3HF_p$:1631 *ms*; $WS3LF_p$: 1632 *ms.*
The factor script type had a significant effect on the latencies so that words with pointed short vowels were categorized more slowly than words with un-pointed short vowels (all *ps* <0.04). Thus, adding short vowels to the script can break the automatic reading process in Persian readers (see (Taha & Azaizah-Seh, 2017) for the same finding and interpretation).

Shimron & Navon, 1982)) showing that spelling transparency had no influence on the performance in lexical decision or semantic categorization of written words in Persian, Arabic, or Hebrew. The conflict between the results of the semantic categorization experiment in this dissertation and the recall experiment by Baluch and Dana (2006) may be explained by methodological differences. The recall experiment was performed after a written word naming task. Participants were first asked to name the written words, and after a short break, they were asked to recall as many words as they could from the naming experiment. The participants were aware of the upcoming recall task in advance. The results of word naming experiments in Persian show that spelling transparency has an influence on the accuracy and latency of naming, such that opaque written words are named more slowly and less accurately than transparent written words (e.g., (Rahbari, 2019; Bakhtiar & Weekes, 2015)). However, Baluch and Dana (2006) did not find any difference between naming latencies for opaque and transparent written words in the naming task administered prior to the recall task. It could, thus, be the case that in the naming experiment by Baluch and Dana, the participants prepared the naming responses to transparent words more quickly than to opaque words, but used some of the response time to rehearse or keep the stimuli in memory. Conversely, for opaque stimuli, more time was spent on naming and little or no time was left to rehearse or keep the opaque stimuli in memory. Therefore, the recall rate was lower for the opaque words as compared to the transparent words.

Thus, in experiments in which no phonological output is required, as in semantic categorization for written words, spelling transparency seems to play no role. The equal performance for opaque and transparent written words in semantic categorization suggests that the “lexical quality” (Perfetti & Hart, 2002) is the same for transparent and opaque words. The quality of lexical entries is assumed to be equal when lexical entries can be identified from written input, and further processing, i.e., semantic processing, can be performed on identified lexical entries equally efficiently.

6-4. Conclusion

The results of the reading experiments in this dissertation and most of the studies reviewed in the Sections 6-2-1 and 6-2-2 showed that when reading tasks require phonological output (written word naming), responses to opaque written words are less accurate (Rahbari, 2019) and slower (Bakhtiar & Weekes, 2015) than to transparent written words in the conventional writing. However, the latency and accuracy for opaque and transparent written words do not differ in other tasks that do not require phonological output, such as lexical decision (Baluch, 1993) and semantic categorization of written words (our data). Moreover, pointing short vowels facilitates naming (Hosseini Almadani, 2016), but it has no influence on the performance in lexical decision (Koriat, 1984) or on the accuracies of semantic categorization of written words (our data). Thus, spelling transparency is a crucial factor in naming, but not in lexical decision and semantic categorization tasks.

The poor performance of Persian readers for opaque words compared to transparent words in naming tasks, but not in lexical decision or semantic categorization tasks can be explained by the word neighbors that are activated while processing written input, and criteria that are recruited to respond in these tasks (Seidenberg, Waters, & Barnes, 1984). The visual processing of written inputs activates orthographic and phonological neighbors of the input stimuli in the process of word recognition (McClelland & Rumelhart, 1981). In this process, the activation of phonological information lags behind the activation of orthographic neighbors.

In the case of opaque words, the visual processing of written opaque stimuli activates similar lexical orthographic entries that are inconsistent with respect to the phonological information they provide for short vowels. Assume that visual processing of an opaque written input such as <mshk> /m**æ**ʃk/ (pitcher) activates these similar lexical orthographic entries while lexicon is being accessed, for instance, <tshk> /toʃ**æ**k/ (mattrass), <khshk> /x**o**ʃk/ (dry), <msht> /m**o**ʃt/ (fist), <msht> /m**æ**ʃt/ (dense), <mshgh> /m**æ**ʃɣ/ (assignment), <mshki> /m**e**ʃki/ (black), <mshkl> /m**o**ʃkel/ (problem), <rshk> /r**æ**ʃk/ (envy), <kshk> /k**æ**ʃk/ (curd), <mushk> /muʃ**æ**k/ (missile), <?shk>, /?**æ**ʃk/ (tear), etc. Considering the short vowel(s), the phonological information that these lexical orthographic entries provide, is not always consistent. Such inconsistency may cause interference in naming opaque words. Naming opaque words is, thus, influenced by interference from inconsistent phonological information provided by similar lexical orthographic entries with respect to the short vowel(s). Pointing short vowels in the script may change the way the lexicon is accessed from visual access to phonological access, or pointing may help the fast resolution of phonological inconsistency caused by similar lexical orthographic entries.

This is not, however, the case with transparent words. Similar lexical orthographic entries to transparent input words provide almost consistent phonological information with respect to long vowels. Assume that the visual processing of the constituents of a transparent written input <kar> /kɑr/ (work) activates similar lexical orthographic entries in the process of lexical access <kart> /k**ɑ**rt/ (card), <kard> /k**ɑ**rd/ (knife), <b**a**r> /bɑr/ (time), <kakh> /k**ɑx**/ (palace), <mar> /m**ɑ**r/ (snake), <tar> /t**ɑ**r/ (vague), etc. The phonological information provided by these orthographic neighbors is consistent with respect to the long vowel.

According to Seidenberg and associates (Seidenberg, Waters, & Barnes, 1984), phonological processing in the written modality is not something in the control of readers to switch it on for naming and off for lexical decision or semantic categorization tasks. Readers, however, can select the criteria based on which responses are made. Responses in written word naming need to be made based on phonological criteria; but in lexical decision or semantic categorization of written words, responses can be made based on other criteria (e.g., visual criteria) which are processed very fast. If responses can be made based on visual criteria before phonological information becomes available, task performance may not be affected by the spelling transparency; otherwise, if the performance is delayed until phonological information is available, spelling transparency may influence the performance regardless of the type of the task. Therefore, the effect of the spelling transparency depends on the time it takes to respond and the criteria that are used to respond.

The null effect of spelling transparency in the semantic categorization task suggests that responses were made based on visual criteria before phonological information became available. Moreover, it can be explained why spelling transparency influenced the performance in a lexical decision task with low frequency opaque and transparent written words in Persian in the study by Baluch (1993). The process of lexical decision may have been delayed for the low frequency words and the availability of phonological information by the time of response would create a problem in case of any inconsistency in the available phonological information.

To sum up, reading opaque words may be less accurate and slower than reading transparent words when lexical entries with the similar spelling to that of written inputs provide inconsistent phonological information with respect to vowels (e.g., <mshk> /mæʃk/, <khshk> /xoʃk/, <tshk> /toʃæk/), and such inconsistent phonological information is used to respond, or is available by the time of response. Moreover, different effects of spelling transparency in naming tasks and in lexical decision or semantic categorization tasks makes it unlikely that lexical access is necessarily mediated by phonology for written words in an opaque language such as Persian. If lexical access was necessarily mediated by phonology, the effect of spelling transparency should have been observed in all these tasks.

As for naming pseudoword with long vowels or pointed short vowels, the results showed that pseudowords with long vowels could be named more accurately than pseudowords with pointed short vowels. Pseudowords are named by phonological recoding, and phonemic awareness has an influence on phonological recoding (González-Valenzuela, Díaz-Giráldez, & López-Montiel, 2016). In word naming, phonological recoding is performed for consonants and long vowels in the Persian conventional writing by Persian adult readers. However, since short vowels are missing in the conventional writing, phonological recoding is never performed for short vowels in words. This process may have led to lower phonemic awareness for short vowels as compared to long vowels so that even when short vowels are pointed, phonological recoding and assembling sounds to name pseudowords are less accurate in pseudowords with pointed short vowels. This point may indicate the nature of the relationship between phonemic awareness and reading. Phonemic awareness is a crucial factor in the early process of reading acquisition for phonological recoding and assembling sounds to utter words. In turn, when readers become skilled, the reading process affects phonemic awareness. Thus, there is a reciprocal relationship between reading and phonemic awareness (Morais, 1991). In addition to the contribution of phonemic awareness in pseudoword decoding, phonemic awareness also helps retaining and manipulating phonemes in phonological working memory (Baddeley, 2012). Thus, the poor performance in naming pseudowords with pointed short vowels as compared to pseudowords with long vowels can suggest lower phonemic awareness and lower memory span for short vowels as compared to long vowels in Persian.

6-5. Correlational analysis

According to the orthographic effect hypothesis, the coactivation of orthographic information while processing phonological information can reinforce the activation level of the phonological information (Ziegler, Ferrand, & Montant, 2004; Ziegler & Ferrand, 1998). Or phonological information can be restructured in terms of the orthographic information (Taft, 2011). Accordingly, it was hypothesized that spoken stimuli with short vowels could be processed less accurately than the spoken stimuli with long vowels in Persian because short vowels, in contrast to long vowels, are not orthographically represented. Since orthographic effect emerges after learning to read and write; therefore, reading experience and reading proficiency were expected to influence the emergence and the size of the orthographic effect. Illiterates were not expected to show the orthographic effect, and the size of the effect was expected to be larger in literates who were more proficient in reading Persian conventional writing. On the other hand, the effect size was expected to be smaller in literates who were more proficient in reading Persian unconventional writing.

To find whether this was the case, a correlational analysis was performed between the performance in the auditory ABX discrimination task and the performance in the reading tasks for literate participants. To perform the correlation analysis, the performance (error rate) was aggregated for vowel category and lexicality per participant in each task. In this way, each participants had some scores showing her/his performance for containing short vowels and containing long vowels (pseudo)words in that task.

Then, the correlational analysis was performed between the performance for the same vowel category and lexicality (corresponding conditions) across tasks for each participant (see Table 6). The corresponding conditions across the tasks are shown with the same font type in Table 6.

In addition to the cross-task correlations, a correlational analysis was performed within the task. This correlation was performed between the performance for the pointed and un-pointed written words with short vowels in the semantic categorization task (see Table 6). This correlation could show us how proficient/poor readers in the un-pointed writing performed in the pointed writing.

Table 6. Corresponding conditions for the correlational analysis between the performance in the auditory ABX discrimination task and reading tasks (naming written pseudowords and semantic categorization for written words).

Task	Naming written pseudowords	Auditory ABX discrimination	Semantic categorization with written words
Condition	*Pseudowords with short vowel*	*Pseudowords with short vowel*	*Words with short vowel (un-pointed)*
	Pseudowords with long vowel	Pseudowords with long vowel	*Words with short vowel (pointed)*
		Words with short vowel	Words with long vowel
		Words with long vowel	

The results of the Spearman correlation analysis in the SPSS (IBM Corp., 2012) showed no significant correlations between the performances for the corresponding conditions across the tasks (all *ps*> 0.37). However, the correlation was significant between the performance for the pointed and un-pointed written words with short vowel in the semantic categorization task [r=0.6, p=0.001]. Better readers in the pointed writing (unconventional writing) were also better readers in the un-pointed writing (conventional writing) for the written words with short vowel. Poor readers had difficulty in reading the conventional writing; and adding short vowels could not help them improve the reading score.

Therefore, the prediction was not confirmed. The performance in the ABX discrimination task did not correlate with the performance in the reading tasks across the corresponding conditions. Better readers in the conventional writing were not poor discriminators in the ABX task, and better readers in the unconventional writing were not better discriminators in the ABX task. This could show that the underlying cognitive mechanisms could be different across these two tasks, i.e., the phonological processing in the reading tasks may differ with the phonological processing in the ABX discrimination task (Perfetti, Beck, Bell, & Hughes, 1987; Morais, Bertelson, Cary, & Alegria, 1986). The mechanism used to discriminate auditory minimal pairs was different from the mechanism used to decode the printed stimuli into phonemes to name them; or processing the written stimuli to categorize them semantically.

Considering the levels of processing (Morais & Kolinsky, 1995; Karmiloff-Smith, 1995; Morais & Kolinsky, 1994; Bialystok, 1986), Metalinguistic abilities (e.g., reading) need a high level of the analyzed linguistic knowledge and a high level of control over cognitive processes which may not be the relevant factors in discriminating auditory (pseudo)minimal pairs, i.e., linguistic ability.

Lack of any correlation between the performance in the reading and discrimination tasks can also be explained by different strategies deployed to perform these tasks. In reading, the segments must be processed to attain the task goal. The evidence in the ABX discrimination task showed that the participants approached the task by the integral strategy and not by the segmental match strategy. In the integral match strategy, the integral closeness of the minimal pairs in the perceptual space was used to discriminate the auditory minimal pairs (Treiman & Baron, 1981). It was likely to find a significant correlation between the performance in the ABX discrimination and reading tasks, if the participants had approached the ABX discrimination by the segmental match strategy.

Thus, the phonological processing underlying (pseudo)words discrimination and reading could be of different types which led to a dissociation between the performance in the auditory discrimination and reading tasks.

Chapter 7: Auditory primed lexical decision experiment

7-1. Introduction

In the domain of written word recognition, it has been shown that a link exists between orthography and phonology, so that phonological information can facilitate or impede the recognition of written words (Rastle, 2012). In visual lexical decision, 'yes' responses are slower to the written homophones, e.g., 'maid' (homophone twin: 'made') than to written non-homophones, e.g., 'paid' (Ferrand & Grainger, 2003; Rubenstein, Lewis, & Rubenstein, 1971). In the same vein, pseudo-homophones, e.g., KOAT, have slower 'no' replies than non-pseudo-homophones, e.g., GURD (Vanhoy & Van Orden, 2001). Visual masked primed lexical decision tasks have shown that priming a written word target with a written pseudo-homophone, 'KOAT – COAT', facilitates the recognition of the target word, here COAT, relative to the recognition of the target word in an orthographic control condition, 'POAT-COAT' (Rastle & Brysbaert, 2006; Lukatela, Frost, & Turvey, 1998).

All this evidence may show that phonological information has a leading role in printed word recognition (Xu & Perfetti, 1999; Lukatela & Turvey, 1994).The effect of fast phonology in written word recognition has been incorporated in the Bimodal Interactive Activation Model (Diependaele, Ziegler, & Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005), see Section 1-3 for the difference between the DRC and BIAM models in the fast sub-lexical phonological activation).

Studies in the domain of spoken word processing have also found the trace of orthographic knowledge in spoken word processing since the 1930s to date (Olson, 1998). In early studies on the effect of literacy on speech processing, there was no distinction between the effect of schooling and literacy per se (Olson, 1998). Later on, thanks to the development of experimental designs and techniques, studies started using more rigidly controlled variables and focused on the role of literacy per se; that is, the role of orthographic knowledge on spoken word processing (Kolinsky, 2015).

A robust finding is the so-called (feedback) consistency effect: when the link between phonology and orthography is inconsistent, auditory word recognition slows down (Ziegler, Petrova, & Ferrand, 2008). Consistency is a graded property and refers to the nature of mapping sounds to spelling in a language. In consistent cases, there is a one-to-one relationship between sounds and spellings; for example, /b/ and /iŋ/ are consistently spelled by <b> and <ing> in English respectively. This is not the case, however, with /f/ and /i:d/ which are represented by multiple spellings <f/ph> (in 'foam' and 'phone', for example) and <ead/eed> (in 'bead' and 'deed').

In an auditory lexical decision task, Ziegler and colleagues (Ziegler, Petrova, & Ferrand, 2008) found that English and French words with inconsistent onset, inconsistent rhymes, or both were processed more slowly than words with consistent units. For instance, 'toast' was processed more slowly than 'luck', since the rhyme /əʊst/ can be inconsistently spelled as <oast> or <ost> in e.g., 'toast' and 'ghost', but the rhyme in 'luck' /ʌk/ is consistently spelled in English (e.g., in 'luck, buck, suck, duck'). The same finding has been reported in other studies on different languages ((Taft, 2011; Ventura, Kolinsky, Pattamadilok, & Morais, 2008; Chereau, Gaskell, & Dumay, 2007; Perre & Ziegler, 2007; Ziegler, Ferrand, & Montant, 2004; Slowiaczek, Soltano, Wieting, & Bishop, 2003); for a review on this topic, see (Kolinsky, 2015)).

Hence, it has been concluded that orthographic knowledge is automatically triggered in spoken word processing leading to the consistency effect (Ziegler, Ferrand, & Montant, 2004). Based on these experimental findings, Grainger and associates (Grainger, Muneaux, Farioli, & Ziegler, 2005) developed the Bimodal Interactive Activation Model (BIAM) of visual and auditory word recognition to capture the role of orthographic knowledge in spoken word processing, not considered in classic models of spoken word recognition, such as the TRACE (McClelland & Elman, 1986) and Cohort (Marslen-Wilson & Tyler, 1980) models. In BIAM, there are bidirectional lexical and sublexical links between spoken and written codes. Orthographic or phonological information can, furthermore, flow from the lexical to the prelexical level within modality through a top-down link.

Compared to real words, the consistency effect has not robustly observed for pseudowords (Cutler & Davis, 2012). Moreover, it is typically observed in tasks that require some degree of lexical involvement, such as the lexical decision task (Ventura, Morais, & Kolinsky, 2006), but for tasks that do not necessarily require lexical involvement, the results are mixed. For example, Ventura and colleagues (Ventura, Morais, Pattamadilok, & Kolinsky, 2004) did not find a consistency effect in a shadowing task in which French adult readers were asked to repeat auditory items that were spelled consistently or inconsistently. By contrast, a consistency effect was observed in a French shadowing experiment conducted by Ziegler and colleagues (Ziegler, Ferrand, & Montant, 2004). Thus, the automaticity of the involvement of orthographic knowledge resulting in the consistency effect has been questioned by some authors in general (Cutler & Davis, 2012; Cutler, Treiman, & Van Ooijen, 2010) or at least for the prelexical level (Pattamadilok, Kolinsky, Ventura, Radeau, & Morais, 2007; Ventura, Morais, Pattamadilok, & Kolinsky, 2004).

Cutler and associates designed a phoneme goodness rating experiment (Cutler & Davis, 2012) and a phoneme detection experiment (Cutler, Treiman, & Van Ooijen, 2010) to examine the consistency effect and the automaticity of the underlying mechanism in the auditory modality. In the phoneme goodness rating experiment (Cutler & Davis, 2012), participants listened to phonetic tokens of [s] that were embedded in words and pseudowords and varied along a continuum with a canonical /s/ token at the mid-point. The participants were asked to rate which token was the best representative for the target phoneme /s/. The mid-point /s/ token was rated higher in words spelled with <s/ss> (e.g., in 'bless') than in words spelled with <c> (e.g., in 'voice'). This result was taken as an indication for the application of orthographic knowledge in spoken word processing. However, this consistency effect was only found for spoken words, not for spoken pseudowords, i.e., the token ratings did not differ between pseudowords such as 'pless' and 'floice'.

In the auditory phoneme detection task (Cutler, Treiman, & Van Ooijen, 2010), the participants were asked to respond as soon as they could detect pre-specified target phonemes in spoken words. The target phonemes occurred word initially, and they were spelled consistently (/b, m, t/) or inconsistently (/f, s, k/). The carrier words (words that contained the target phoneme) were put into two different contexts in two different lists. In the first list, the carriers were put in a context that included many irregular words in which some graphemes were not pronounced (e.g., 'kneel, wrinkle, tongue', etc.) as fillers (words that did not contain the target phoneme), whereas the context in the second list included many regular words as fillers ('nip, wiggle, lung', etc.). The lists were assigned to two groups of participants. The consistently spelled phonemes /b, m, t/

were detected faster than the inconsistently spelled phonemes /f, s, k/ in the list that contained many words with irregular spellings as fillers. The consistently spelled phonemes were not detected faster than the inconsistently spelled phonemes in the list that contained words with regular spelling as fillers. The authors concluded that the consistency effect was due to a strategic deployment of orthographic knowledge in the context where spelling inconsistency was rendered salient by including many irregular fillers. They argued that if the consistency effect was automatic, it should have been observed for both real words and pseudowords in the phoneme goodness rating experiment due to the partial similarity between real words (e.g., 'bless' and 'voice') and pseudowords (e.g., 'pless' and 'floice'). Moreover, if the consistency effect was automatic, it should have been observed in the phoneme detection experiment regardless of the surrounding items (fillers).

The lack of a consistency effect in the shadowing task by Ventura and Colleagues (Ventura, Morais, Pattamadilok, & Kolinsky, 2004) was taken as additional evidence against the automatic activation of orthographic knowledge in speech processing. In sum, according to Cutler and colleagues, a consistency effect that is limited to the lexical level where the orthographic information is present, and to a context where orthographic information is rendered salient cannot be considered as an automatic activation of orthographic knowledge in speech processing. The evidence in these experiments rather suggests that orthographic information is taken into account by a decision-making mechanism at a post-lexical level where the attention was drawn toward spelling inconsistency (Cutler & Davis, 2012; Cutler, Treiman, & Van Ooijen, 2010). Such a decision-making mechanism, taking sublexical, lexical, and contextual information into account, is assumed in the Merge model (Norris, McQueen, & Cutler, 2000).

Depending on task demands, the decision-making mechanism will put more weight on the information arriving from the bottom (sublexical information), from the top (lexical information), or from the context. The higher the task demand, the higher the sensitivity of the decision-making mechanism to higher-level knowledge.

Thus, there are two main opposing accounts regarding the influence of orthographic knowledge on speech processing. The BIAM (Diependaele, Ziegler, &

Grainger, 2010; Grainger, Muneaux, Farioli, & Ziegler, 2005) predicts the automatic activation of orthographic knowledge, whereas the Merge model predicts a strategic employment of orthographic knowledge while processing speech. Yet, other models consider no role for orthographic knowledge, neither automatic nor strategic, while processing speech (see (Weber & Scharenborg, 2012) for different models of speech processing). In short, there is no consensus about the role of orthographic knowledge and its nature in speech processing.

The present study aimed to contribute to the cross-linguistic literature on this topic. To investigate whether orthographic knowledge was automatically activated while processing spoken words in Persian, we administered an auditory primed lexical decision task with Persian literate participants. The findings of other experiments with the same paradigm using English stimuli (Slowiaczek, Soltano, Wieting, & Bishop, 2003; Jakimik, Cole, & Rudnicky, 1985) showed a facilitatory effect for processing targets ('mess') sharing the initial sounds and spelling with carrier primes ('message'). The same facilitation was not found when the targets shared either the sounds or the spelling with the carrier primes, as in 'definite-deaf' or 'legislation-leg'. Slowiaczek and colleagues (Slowiaczek, Soltano, Wieting, & Bishop, 2003) found the same results with the same paradigm. They found a consistency effect for both words and pseudowords while masking the relationship between the primes and the targets by including a large number of unrelated prime-target pairs. Therefore, they interpreted their results as a robust indication for an automatic effect of orthographic knowledge in a purely auditory task.

Studies investigating the orthographic consistency effect on speech processing have mainly been conducted in English, French, or Portuguese. The present investigation was performed on Persian to examine the orthographic consistency effect in a language that (to the best of author's knowledge) has not been studied with respect to this topic by the time of administering the experiment. Cross-linguistic studies are important, because the evidence shows that the consistency effect is related to reading and writing processes (Saletta, Goffman, & Hogan, 2016; Pattamadilok, Nelis, & Kolinsky, 2013), and these processes differ across different orthographic systems (Frost, 2005; Katz & Frost, 1992). For example, Pattamadilok and associates (Pattamadilok, Morais, De Vylder, Ventura, & Kolinsky, 2009) found that the emergence and the development of the consistency effect differs between French and Portuguese, because these two languages differ with respect to the consistency of the links between sound and spelling, with French being more inconsistent than Portuguese.

In Persian, some phonemes are arbitrarily spelled in different ways. For example, /t/ is represented by either of the graphemes ت or ط in words such as تِن /tæn/ 'body' and طِرح /tærh/ 'design' (Note that Persian is written from right to left.). Such phonemes and words are called inconsistent phonemes and words in this dissertation, as they have multiple orthographic representations. In Persian, not only /t/ but also the phonemes /?, s, h, z, ɣ/[38] are spelled in several ways. In contrast, phonemes that are always spelled with the same grapheme are called consistent phonemes. For example, /b/ is spelled consistently with ب in Persian across words. In this regard, the sound-to-spelling mapping is inconsistent, whereas the spelling-to-sound mapping is consistent in Persian: either of the graphemes ت or ط always represents the same sound /t/. Hence in Persian, the effect of the pure sound-to-spelling consistency can be examined.

In our primed auditory lexical decision task, we hypothesized that when corresponding primes and targets share the same initial sound and spelling (e.g., **کِش** – کِشور /keʃvær-**keʃ**/ 'country-elastic band'), lexical decision latencies for such targets should be shorter than when the same targets are primed by unrelated word forms (e.g., **کش** – لگن /lægæn-**keʃ**/ 'basin-elastic band'). Whereby contrast, when corresponding primes and targets share the same initial sound, but with different spellings (e.g., **تِن** - طِناب /tænɑb-**tæn**/ 'rope-body'), lexical decision latencies for these inconsistent targets should not be faster compared to the same targets primed by unrelated word forms (e.g., **تن** - کوچه /kuʧeh-**tæn**/ 'alley-body'). The mechanism involving in the consistency effect can be explained as follows. Based on the Cohort model for spoken word recognition (Marslen-Wilson & Tyler, 1980), listening to a spoken input activates word candidates (cohorts) that overlap with the spoken input in initial sounds. The residual activation of word candidates following the recognition of the spoken input leads to the faster

[38] <?, ا; س, ص, ث; ه, ح; ز, ض, ذ, ظ; ق, غ>

recognition of those word candidates (Pisoni, Nusbaum, Luce, & Slowiaczek, 1985). Therefore, targets should be expectedly recognized faster after related primes (/keʃvær-**keʃ**/, /tænɑb-**tæn**/) than after unrelated primes (/lægæn-**keʃ**/, /kuʧeh-**tæn**/). Additionally, orthographic knowledge may reorganize the structure of cohorts in a way that word candidates with similar and dissimilar spellings are in different phonological networks (Muneaux & Ziegler, 2004). The more the overlap between the spoken input and word candidates (i.e., phonological as well as orthographic overlaps), the higher the level of activation for word candidates while listening to the spoken input, and the higher the residual activation of word candidates after the recognition of the spoken input. In this case, the residual activation of /**keʃ**/ کِش after listening to /keʃvær/ کِشور would be higher than the residual activation of /**tæn**/ تِن after listening to /tænɑb/ طِناب leading to faster recognition of /**keʃ**/ کش due to more overlap in /keʃvær-**keʃ**/ as compared to /tænɑb-**tæn**/.

The hypothesized outcome would show that orthographic knowledge is deployed in a completely auditory task, and that inconsistency of spelling between corresponding primes and targets modulates the facilitatory effect of phonological priming[39]. Moreover, we expected this pattern of results to be observed in any context, i.e., whether priming by unrelated words precedes or follows priming by phonologically related words. This outcome would be an indication for a robust consistency effect. In other words, we considered the necessary counterbalancing of the experimental conditions as a context manipulation to which an automatic consistency effect should be resistant. Inconsistent orthographic codes generated while listening to inconsistent targets in related trials were expected to be resolved faster when unrelated trials preceded related trials than when related trials preceded unrelated trials.

[39] For the facilitatory effect of phonological priming on target word recognition, see (Dufour & Peerman, 2004; Slowiaczek, Nusbaum, & Pisoni, 1985).

7-2. Experiment

7-2-1. Participants

Twenty monolingual Persian literates (8 females) accepted the invitation of the experimenter to participate in this experiment. All were living in Iran, Fars province, and were aged between 20 to 50 years (*SD*=7.5). They were master students or postgraduates active in academic fields not related to language or linguistics. All of them were exposed to the Persian conventional writing for at least fourteen years, and reported daily reading in Persian. They were right-handed without any reported hearing or motor problems. The participants were paid the equivalence of four Euros in the local currency, and signed the consent form for their participation and for the anonymous publication of the results.

7-2-2. Stimuli

Forty monosyllabic Persian real words (all nouns) were selected as targets, half of which were consistent words, and the other half were inconsistent words. The initial phonemes of the consistent words were always spelled by the same graphemes, whereas the initial phonemes of the inconsistent words could be spelled in several ways. For example, the initial phoneme of کش /keʃ/ 'elastic band' is spelled consistently with <ک> across words, whereas the initial phoneme of تن /tæn/ 'body' can be spelled inconsistently with <ت> or <ط> across words.

The consistent target words were primed by twenty disyllabic carrier words whose initial parts overlapped with the targets in both sound and spelling, such as کشور **- کش** /keʃvær - **keʃ**/ 'country - elastic band'. The inconsistent target words were primed by twenty disyllabic carrier words whose initial parts overlapped with the targets in sound, but not in spelling, such as طناب **- تن** /tænɑb – **tæn**/ 'rope - body'. Thus, in both conditions corresponding primes and targets were phonologically related, but the two conditions differed with respect to the spelling consistency of the initial consonants. Henceforth, these two conditions will be called 'related-consistent' and 'related-inconsistent' respectively.

In two further conditions, the same targets were preceded by forty disyllabic phonologically unrelated word forms e.g., لگن -**کش** /lægæn-**keʃ**/ 'basin - elastic band' and کوچه - **تن** /kuʧeh-**tæn**/ 'alley – body', henceforth called 'unrelated$_{Cons}$' and 'unrelated$_{Incons}$' or briefly 'unrelated'. Considering the two factors 'phonological relatedness' (related/unrelated) and 'orthographic consistency' (consistent/inconsistent) there were four experimental conditions:

'related-consistent': /keʃvær - **keʃ**/ کشور -**کش**

'related-inconsistent': /tænɑb - **tæn**/ طناب - **تن**

'unrelated$_{Cons}$': /lægæn - **keʃ/** لگن - **کش**

'unrelated$_{Incons}$': /kuʧeh - **tæn**/ کوچه - **تن**

The subscripts in the unrelated conditions are used to identify the corresponding target type which followed the unrelated primes in those trials. For example, unrelated$_{Cons}$ showed that the target that followed the unrelated prime was an orthographically consistent target.

To illustrate with corresponding English examples, 'napkin – nap' would be a prime –target pair of the 'related-consistent' condition, and 'phantasy – fan' of the 'related-inconsistent' condition. The same targets 'nap' and 'fan' would be preceded by the unrelated primes 'ruler – nap' and 'album- fan' in the 'unrelated$_{Cons}$' and 'unrelated$_{Incons}$' conditions respectively. In our experiment, the conditions were matched for these features:

- In the phonologically related conditions, the targets were the first syllables of the primes in one half of the trials of that condition /ser.væt – ser/; and for the other half, the targets were the first syllable plus the onset of the second syllable of the primes /ʤɑ.med - ʤɑm/.
- In half of the trials of the 'related-inconsistent' condition, the targets had the dominant spelling, and in the other half of the trials of the same condition, the targets had the subordinate spelling. The dominancy was defined in terms of the most frequent spelling (letter) used to represent an inconsistently spelled phoneme. Five native Persian speakers rated the frequency of the spellings (for example which of the letters ت or ط is more frequently used to represent the sound /t/). It should be noted that there is an arbitrary relationship between

inconsistent phonemes and their orthographic realizations. That /t/ is realized as ت in تن and as ط in طناب is an arbitrary but conventional matter, i.e., it is not rule-governed.

- Each target was used once in the related condition (e.g., طناب - **تن** /t̲ænɑb-**t̲æn**/) and once in the unrelated condition (e.g., کوچه - **تن** /kuʧeh – **t̲æn/**). There were 50 to 83 intervening trials between a target and its repetition. This number varied randomly from target to target.
- The order of the presentation of a given target with the related prime and with the unrelated prime was balanced within lists and counterbalanced across lists. Within a list, half of the inconsistent and half of the consistent targets were primed first by the unrelated words and then by the related words. For the other half, the order was reversed.
- The uniqueness point of the targets was always on the third phoneme.
- The word frequencies of the target words and their phonological neighborhood sizes and frequencies were obtained from a Persian word frequency database (Esfahbod, 2012) and the Mo'in dictionary (Mo'in & Shahidi, 1972), respectively.
- Primes and targets were selected in a way that no semantic relationship was between corresponding primes and targets.

The conditions, examples and their features are presented in Table 7. The critical phonemes and letters are underlined.

Additionally, eighty disyllabic real words and eighty monosyllabic pseudowords were used as the filler primes and targets (e.g., مدار - **نپ** /mædɑr–**nop**/ 'circuit'). The real word primes were different in the experimental and the filler trials. The pseudowords were created by replacing the last phonemes of three-phonemic real words (which were different from the targets in the experimental conditions) to keep the uniqueness point on the third phoneme of the targets across the real word (experimental targets) and pseudoword (filler) targets.

Table 7. Characteristics of the stimuli in the primed lexical decision experiment.

Orthography	Consistent		Inconsistent	
Phonology	**Related**	**Unrelated**	**Related**	**Unrelated**
Example (prime-**target**)	keʃvær - **keʃ**	lægæn - **keʃ**	tænɑb - **tæn**	kuʧeh - **tæn**
Meaning	country-elastic band	Basin-elastic band	rope- body	alley-body
	کش - کشور	**کش** - لگن	**تن** - طناب	**تن** - کوچه
Spelling	KSHVR-**KSH**	LGN-**KSH**	TNAB-**T°N**	KUCHH-**T°N**
Mean frequency - prime	2301	1693	2383	1599
Mean frequency - target	3920	-	3779	-
Mean number of phonemes - prime	5.5	5.1	5.45	5.05
Mean number of phonemes - target	3	-	3	-
Mean number of phonological Neighbors - Target	15.1	-	14.75	-
Mean log Frequency for Phonological Neighbors - Target	3.6	-	3.65	-
Mean duration of targets (second)	0.5	-	0.52	-

To reduce the chance of anticipating targets upon hearing the primes, nineteen percent of filler primes contained initially embedded real words that could have been used as related targets, but weren't. For example, in the prime-target filler pair /mædɑr-nop/, the initial part of /mædɑr/ is a real word (i.e., /mæd/ 'ebb') in Persian which could have been used as the target in that trial (/mædɑr-mæd/), but instead the target in that trial was the pseudoword /nop/. The performance for filler items was not included in the statistical analysis.

The items were randomized in four different orders across four lists. Ten warm-up trials were added to the beginning of each list, and each list was assigned to five participants. Moreover, the randomization within a list was such that trials from the same condition did not immediately follow each other, and expected successive key presses on the same key were limited to three [(i)].

7-2-3. Procedure

Following Meyer and Schvaneveldt (1971), in this auditory primed lexical decision experiment, an auditory target word was preceded by an auditory prime word that shared or did not share some of its initial phonemes. The task instruction was orally explained to each participant before the experiment by the experimenter. The participants were asked to listen to the spoken items played successively in each trial (i.e., prime - target), and decide whether the second item in each trial was a real word or not. It was explained to them before the experiment that some items were pseudowords (possible but non-existing word in Persian). They were asked to press the response key as fast as and as accurately as possible. The written instruction was also displayed at the beginning of the experiment on the laptop screen, and the participants were asked to read it carefully and ask if something was unclear.

The experiment started by pressing the space button by the participant. Before each trial, a calm beep sound was played to signal the new trial. 200 *ms* after the beep sound, the prime and then the target were spoken successively. The inter-stimulus interval was set on 500 *ms* within the trials. The numbers '1' and '2' were displayed respectively on the monitor screen when the spoken prime (1) and target (2) were presented. From the onset of the target (second item in each trial), participants had 2000 *ms* time to decide and reply by pressing the response keys whether the second item (target) was a real word or not. The inter-trial interval was 1000 *ms*.

The spoken stimuli were loaded into Psychopy V.2020.1.3 (Peirce, et al., 2019) installed on an Asus laptop Vivo Book. The left and right arrow keys on the keyboard of the laptop were used as the response keys for 'yes' and 'no' replies (right key: yes, it is a real word; left key: no, it is not a real word). The participants used their dominant hand for 'yes' replies and listened to the stimuli through a Sony headphone (WH-100XM3).

After finishing the auditory primed lexical decision task, the participants were asked to take part in a spelling post-test after a short break. In the post-test, they were asked to listen to the spoken words which were played one by one, and write them down on paper. This task was intended to make sure that the participants knew the correct spellings of the inconsistent primes and targets used in the auditory primed lexical decision task. For that purpose, the inconsistent primes and targets were randomized such that each trial contained one word and successive trials did not start with the same sound. A new trial (word) was signaled by a beep tone. 200 *ms* after the beep sound, a word was played, and from the offset of the spoken words, the participants had 5000 *ms* time to write them down on paper. The inter-trial interval was 1000 *ms*.

7-3. Data analysis

Incorrect replies (9.5%), not-replied items (5.4%), trials with reaction times outside the cutoff value (±3SD from the mean reaction time in each condition; 0.6%), warm-ups, and fillers were discarded from the reaction time analysis.

To find whether orthographic consistency had an influence on the lexical decision latencies for target words, phonological relatedness (related vs. unrelated) and orthographic consistency (consistent vs. inconsistent) between primes and targets within trials were manipulated (see Figure 20). The lexical decision latencies for consistent targets preceded by phonologically related words were expected to be faster than the lexical decision latencies for the same consistent targets preceded by phonologically unrelated words. But the same facilitation was not expected for inconsistent targets when preceded by phonologically related or unrelated words.

Linear mixed effect regression analyses, lmer model (lme4 package (Bates, Mächler, Bolker, & Walker, 2015)) fitted with lmerTest and REML packages (Kuznetsova, Brockhoff, & Christensen, 2017) were used in R (R Core Team, 2014) to analyze the reaction times. The effect of each factor was computed by comparing the model including all factors with the model without that target factor using ANOVA tests.

The values for degree of freedom (Satterthwaite method) and F values were obtained using ANOVA tests in R. The post hoc tests were run using the lsmeans function in R (Russell, 2016). Intercept for 'participants' and 'items' were inserted into all models as random factors.

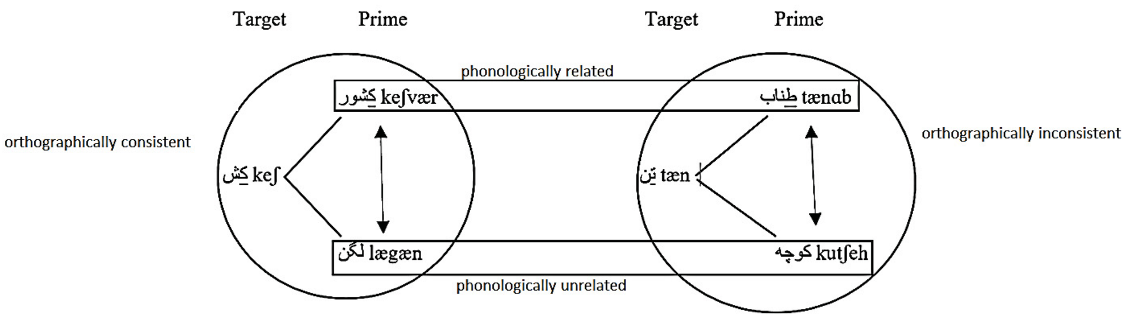


Figure 20. Factors, groupings, and examples in the primed lexical decision experiment. The arrows show the priming effect.

The grouping for the corresponding conditions is illustrated schematically in Figure 20. The arrows within the ellipses represent the corresponding conditions for the priming effects.

7-4. Results

7-4-1. Consistency effect

Inconsistent targets with dominant spelling were recognized as fast as inconsistent targets with subordinate spelling (t (22.6) =-0.03, p=0.8). Moreover, the syllabic overlap between the corresponding primes and targets did not have any influence on the recognition time for the targets (*t* (34) =0.4, *p*=0.5). The interaction between syllabic overlap and orthographic consistency was also insignificant (*t*(35) =-1.24, *p*=0.2). Thus, the data were collapsed for spelling dominancy and syllabic overlap.

The graph in Figure 21 shows the mean lexical decision latencies of the correct replies for words as a function of phonological relatedness and orthographic consistency. The mean lexical decision latency for consistent targets was numerically faster when they were preceded by phonologically related words (*M*=1137 *ms*, *SD*=41.8) than when they were preceded by unrelated word forms (*M*=1182 ms, *SD*=41.8), resulting in a priming effect of 45 *ms*.

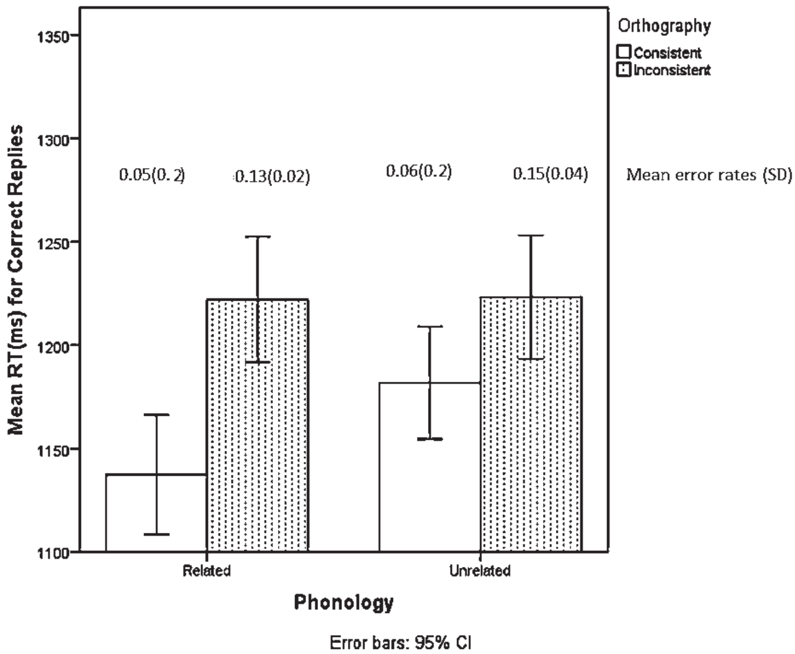


Figure 21. Auditory primed lexical decision by Persian literates. Mean reaction time (RT) in millisecond (ms) for correct replies to words as the function of phonological relatedness and orthographic consistency. The error rates and standard deviations (SD) for each condition are included on the top of the bars.

The mean lexical decision latency for inconsistent targets was the same when they were preceded by phonologically related words (*M*=1225 *ms*, *SD*=42) as when they were preceded or by unrelated word forms (*M*=1225 *ms*, *SD*=42), resulting in a priming effect of 0 *ms*.

The result of the linear mixed effect regression analysis showed significant main effects of orthographic consistency [*F* (1) = 7.23, *p*=0.01] and phonological relatedness

[F (1) =5.82, p=0.016] on the lexical decision latencies for the target words. The interaction between orthographic consistency and phonological relatedness was marginally significant [F (1) = 2.78, p=0.09].

Pairwise comparisons revealed that the mean lexical decision latency for the consistent targets was significantly faster when they were preceded by phonologically related words than when they were preceded by unrelated word forms (t(1299)= -2.97, p=0.003). However, the mean lexical decision latencies for the inconsistent targets did not differ significantly between inconsistent targets that were preceded by phonologically related words and the same inconsistent targets preceded by unrelated word forms (t(1299) = -0.5, p=0.61). The mean lexical decision latency for the consistent targets did not differ significantly from that for the inconsistent targets when both were primed by unrelated word forms (t(51.5) = -1.8, p=0.09). However, the mean lexical decision latency for the consistent targets was faster than the mean lexical decision latency for the inconsistent targets when both were primed by phonologically related words (t(50)= -3.14, p=0.003).

7-4-2. Context effect

As explained earlier, each target was once preceded by a phonologically related word and once by an unrelated word form. Half of the consistent and half of the inconsistent targets were preceded first by phonologically related words and then by unrelated word forms. For the other half, the order of the presentation was reversed. In this part, the results are analyzed in terms of the order of the presentation of the targets (context effect). This analysis allows us to find whether the context had any influence on the lexical decision latencies. In the previous section, the results were interpreted as a function of phonological relatedness and orthographic consistency. Here, the order of the presentation of the targets is also taken into account to find whether context modulated the reaction times across conditions.

Figure 22 shows the mean reaction times for correct replies to words as a function of phonological relatedness and orthographic consistency for the first and second presentations of the targets. The mean lexical decision latency for consistent targets was numerically faster when they were preceded by phonologically related words (first presentation: *M*=1157, *SD*=43.7; second presentation: *M*=1118, *SD*=43.7) than when they were preceded by unrelated word forms (first presentation: *M*=1198, *SD*=43.8; second presentation: *M*=1166, *SD*=43.6), irrespective of presentation order. The mean lexical decision latency for first time presentations of inconsistent targets was numerically slower when they were preceded by phonologically related words (*M*=1282, *SD*=44) than when they were preceded by unrelated word forms (*M*=1255, *SD*=44.3). For second time presentations, the mean lexical decision time for inconsistent targets was numerically faster when they were preceded by phonologically related words (*M*=1168, *SD*=43.8) than when they were preceded by unrelated word forms (*M*=1195, *SD*=44).

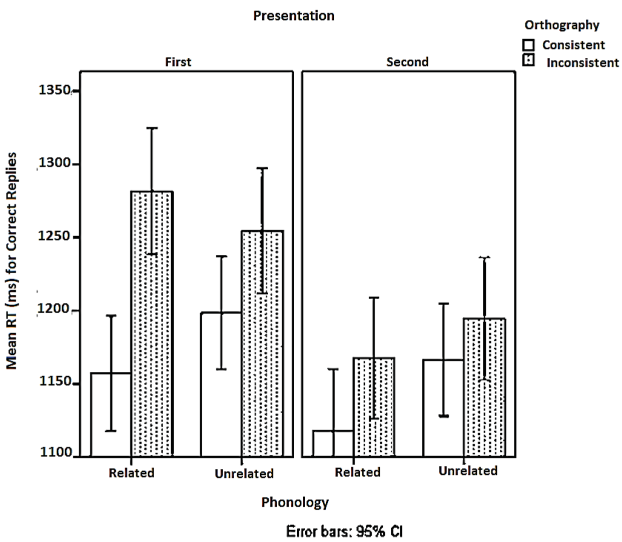


Figure 22. Auditory primed lexical decision by Persian literates. Mean reaction time (RT) in millisecond (ms) for correct replies to words as the function of phonological relatedness and orthographic consistency for the first and second presentations of the targets.

The results of a linear mixed effect regression with the factors ‘phonological relatedness’ (related, unrelated), ‘orthographic consistency’ (consistent, inconsistent), and ‘presentation’ (first, second) showed significant main effects of all three fixed factors on the lexical decision latencies: [‘orthographic consistency’: $F(1)$=7.6, p=0.009; ‘phonological relatedness’: $F(1)$=6.1, p=0.014; ‘presentation’: $F(1)$=44.6, p<0.0001]. Among the two-way interactions, the interaction between ‘presentation’ and ‘orthographic consistency’ was significant [$F(1)$=3.83, p=0.05]. The interaction between ‘orthographic consistency’ and ‘phonological relatedness’ was marginally insignificant [$F(2)$ =3.07, p=0.08]. The interaction between ‘presentation’ and ‘phonological relatedness’ was also insignificant [$F(1) = 1.68$, $p=0.2$]. The three-way interaction between the fixed factors was insignificant as well [$F(1) = 0.92$, $p=0.34$].

Pairwise comparisons revealed that lexical decisions for the consistent targets preceded by phonologically related words took significantly less time than lexical decision for the same targets preceded by unrelated word forms for the first ($t(1296)=-1.98$, $p=0.048$) and the second ($t(1294) = -2.37$, $p=0.018$) presentations of the targets. Yet, lexical decisions for the inconsistent targets preceded by phonologically related words did not differ statistically from the lexical decisions for the same inconsistent targets preceded by unrelated words in the first ($t(1295) = 0.74$, $p= 0.46$) as well as the second ($t(1298) = -1.48$, $p=0.14$) presentations of the targets.

The spelling post-test results showed an overall accuracy of 93%. Discarding the reaction time data of the three least accurate items in the spelling test (/ɣor, ɣod, ?ɑr/) did not change the results of the linear mixed regression analysis reported above.

7-5. Discussion

In an auditory primed lexical decision experiment, we manipulated the orthographic consistency of target words and the phonological relatedness between prime and target words to see whether orthographic consistency would modulate the standard phonological priming effect. Our results showed that this was indeed the case. When consistent target words were preceded by phonologically related words, the lexical decision latency for the targets was on average 45 *ms* shorter than when the same targets were preceded by unrelated word forms. This phonological priming effect was not observed for inconsistent targets: Lexical decision times for such targets did not differ between a condition in which they were preceded by phonologically related words and a condition in which they were preceded by unrelated word forms. Hence, our results show an orthographic effect in a purely auditory task.

This effect could either be due to the orthographic consistency of the target words as such or to the orthographic overlap between primes and targets as consistent target were spelled like the primes whereas inconsistent targets were not. In the former case, lexical decision times for inconsistent target words should have been longer than for consistent target words also in the conditions where both were preceded by unrelated word forms. There was, however, statistically no difference between the decision latencies for the two types of target words in these conditions, suggesting that the absence of a phonological priming effect for inconsistent targets was rather due to the difference in spelling between these targets and their phonologically related primes than due to the inconsistency of the targets as such.

The absence of a phonological priming effect for orthographically dissimilar prime-target pairs could in principle be due to an independent orthographic inhibition effect counteracting the phonological priming effect or to an absence of phonological priming in the case of orthographic dissimilarity. Apart from the fact that an independent orthographic inhibition effect of almost exactly the same size as the phonological priming effect seems rather unlikely, there is also evidence from the literature that favors the second account. It has been suggested that words sharing phonology and orthography (phonographic neighbors) form a network in the lexicon (Siew & Vitevitch, 2019) and there is experimental evidence supporting this assumption. Muneaux and Ziegler (2004), for example, showed that when participants were asked to generate words which ‘sounded’ similar to target words, they tended to name words that were orthographically as well as phonologically similar to the target words. For example being asked to produce a word that rhymed with ‘wipe’, participants tended to produce ‘ripe’ or ‘pipe’ rather than ‘type’. Transferred to the present study, this might mean that coming across the prime word /tænɑb/ which is a Persian real word with the spelling طناب, lexical items would be activated that shared the same sound and spelling with the prime e.g., /tænz/ طنز (joke), but not, or much less so, lexical items that only shared the sound with the prime, e.g., /tæn/تن.

According to the Cohort model (Marslen-Wilson & Tyler, 1980) and the Cohort activation model (Slowiaczek, Nusbaum, & Pisoni, 1985)[40], initially embedded targets are activated when processing the embedding prime word in the phonologically related prime-target sets. Finding a phonological priming effect when primes and targets share

[40] Activation of word candidates based on word initial acoustic-phonetic information during word recognition (Slowiaczek, Nusbaum, & Pisoni, 1985).

their spelling and failure to find a phonological priming effect when primes and targets have dissimilar spelling might mean that cohorts based on shared initial phonology and orthography are processed faster than cohorts based on shared phonology alone. It might be that literacy helps developing lexical networks in which words with similar phonology and orthography are linked, such as pipe, ripe, swipe, tripe, etc. (Muneaux & Ziegler, 2004).

In order to assess whether the observed orthographic effect on phonological priming was due to an automatic orthographic effect or due to a strategic employment of orthographic knowledge, we reasoned that the counterbalancing of the order of related and unrelated conditions can be seen as a kind of context manipulation. We had expected that priming inconsistent targets first by unrelated words and then by phonologically related words would faster resolve spelling inconsistencies activated by inconsistent targets in the related trial, whereas such a fast resolution of spelling inconsistencies in related trials was not expected when inconsistent targets were first primed by phonologically related words and then by unrelated words. Insofar as the orthographic effect would not be influenced by presentation order, it could be considered as being due to a context-insensitive and hence automatic effect. Conversely, an influence of presentation order would point to a context-sensitive, strategic effect. The evidence is in favor of the context-insensitive and automatic orthographic effect: lexical decision latencies for consistent targets were faster in related trials than in unrelated trials, irrespective of whether the unrelated trials preceded or followed the related trials. Likewise, lexical decision latencies for inconsistent targets were equally fast in the related and the unrelated trials, irrespective of whether the unrelated trials preceded or followed the related trials.

Apart from the main findings, the statistics showed a marginally significant interaction between orthographic consistency and presentation order, indicating that the speed-up of lexical decision latencies from the first to the second presentation of the targets was more pronounced for inconsistent targets than it was for consistent targets. This difference might simply be due to a floor effect: lexical decision latencies for consistent targets were already quite fast upon first presentation so that there was little room for even faster processing of these targets.

In general, the findings suggest that orthographic knowledge can restructure the organization of phonological cohorts. The phonological cohort for a spoken word can be ordered in terms of spelling. In cohorts organized like this, friends, i.e., words sharing similar sounds and spellings are closer than enemies, i.e., words sharing similar sounds with different spellings (see Chapter two, Figure 9).

End Notes:

(i) For this experiment, the author used a lexical database available as a WORD file in the Web (Esfahbod, 2012). Written word frequencies from this source seemed to be more dependable than those reported by Google, since the frequencies were generated from texts, such as newspaper articles, that could be assumed to be read by typical Persian speakers in their daily lives. The author generated phonological and orthographic neighbors manually by using Persian dictionaries. This was a very time-consuming process that could be performed for the last experiment only. Accessing n-gram frequencies was operationally impossible. Considering that Persian short vowels are not written, and all the available databases are for written words, spoken n-gram frequencies could not be manually extracted for Persian words with short vowels.

Chapter 8: Summary and general discussion

8-1. Summary of the results

It has been shown that frequent associations are created between phonology and orthography through reading and writing (Van Orden, 1990). Such associations can link the systems involved in processing spoken and written words (Ehri, 1985). These associations can lead to the coactivation of cross-system codes (orthographic codes) when phonological codes are activated (Grainger, Muneaux, Farioli, & Ziegler, 2005), or these associations can restructure phonological codes in terms of orthographic features (Taft, 2011; Pattamadilok, Knierim, Kawabata Duncan, & Devlin, 2010). The empirical evidence shows that orthographic knowledge can shape the perception of speech (Ziegler & Ferrand, 1998). Where the association between phonology and orthography is incomplete (Van Ooijen, Cutler, & Norris, 1992) or inconsistent (Ziegler, Ferrand, & Montant, 2004), speech processing is less accurate or slower.

Such influences of orthography on speech processing have been widely investigated in transparent languages. Most of the investigations found the effect of orthography on speech processing (see Chapter 2). It could be argued that the influence of orthography on speech processing was locked to the nature of the associations between phonology and orthography. The frequent one-to-one correspondences or associates between orthographic and phonological codes in transparent languages could lead to the systems binding (Hebb, 1949), in such a way that the system involved in processing spoken words binds to the system involved in processing written words. Such binding might facilitate or modulate information processing (Spence & Deroy, 2013; van Atteveldt, Formisano, Goebel, & Blomert, 2004) in transparent languages when associations are consistent or inconsistent. However, the case could be different in opaque languages.

The associations between phonology and orthography are not one-to-one in opaque languages. One phonological unit may not have any corresponding orthographic unit, or one phonological unit may be represented by multiple orthographic units. It could, thus, be claimed that the system-binding might not be consolidated in opaque languages leading to weaker or absent effects of orthography on speech processing. Conversely, stronger orthographic effects could be expected in opaque languages as compared to transparent languages due to restructuring of phonological codes in terms of orthographic features that are highly inconsistent in opaque orthographies. Thus, data from different languages with different mapping patterns between phonology and orthography could make it clear whether effects of orthography on speech processing depend on the mapping patterns between phonology and orthography.

To this purpose, I investigated the effects of orthography on speech processing in an opaque language such as Persian. Short vowels, in contrast to long vowels, have no orthographic realizations in the conventional writing in Persian. Moreover, some consonants are polygraphic, while others are monographic in Persian. These two cases were used to investigate the influence of orthography on speech processing in Persian to add more empirical data to the related literature. If speech processing was affected by orthographic features, after a long exposure to Persian conventional writing system, Persian adult literates could be expected to process spoken items with long vowels more accurately than spoken items with short vowels. Furthermore, spoken items including consonants that are spelled in multiple ways could be expected to be processed more slowly than spoken items including consonants that are spelled in one way. Different experiments were conducted, each of which tapped into a certain level of processing to find whether orthography influences speech processing at that level.

The performance in the Persian auditory AX discrimination experiment (Chapter 3) showed that Persian minimal pairs with short vowels were discriminated as accurately as minimal pairs with long vowels by Persian literates and the control groups (German literates and Persian illiterates). Also, words with short vowels could be discriminated well from words with long vowels when the short and long vowels were exceptionally represented by the same grapheme. Items in this exceptional condition were discriminated as accurately as items with long vowels in which the vowels were distinctly spelled by the respective graphemes. Thus, when the mapping pattern between phonology and orthography was one-to-none (in the case of short vowels), two-to-one (in the case of the two vowels that were exceptionally represented by the same grapheme), or one-to-one (in the case of long vowels), the discrimination rate did not differ when the task was performed by Persian literates and the two control groups. Hence, speech discrimination was not affected by orthographic features in the case of vowels in the Persian auditory AX discrimination task performed by Persian literates.

The performance of German literates was, however, affected by orthographic features in the German AX discrimination task (Chapter 3)[41]. Spoken minimal pairs different in vowel length (length is phonemic in German) were poorly discriminated when they had the same spelling. By contrast, spoken minimal pairs differing in vowel length were well discriminated when they had different spellings. All other phonemic contrasts were well discriminated by German literates in this task. The finding that orthographic knowledge could influence the discrimination of German spoken minimal pairs was unpredicted. This effect of orthography was found for real words, but not for pseudowords in the condition where pseudo minimal pairs with potential identical spelling were to be discriminated. That is, pseudo minimal pairs differing in vowel length could be well discriminated when they could have the same spelling considering the spelling-to-sound correspondences in German, showing that the effect of orthography on speech processing was limited to words, and sublexical units were immune to the orthographic effect. This could be an indication that information does not flow top-down as predicted by interactive models of speech processing (McClelland, Mirman, & Holt, 2006; Grainger, Muneaux, Farioli, & Ziegler, 2005), as lexical knowledge was not feed-backed to lower levels. If it were, the effect of orthography should have been observed for both real words and pseudowords due to partial similarity between words and pseudowords. Thus, speech discrimination was affected by German orthographic features in the German AX discrimination task performed by German literates.

Compared to the AX discrimination task, the auditory ABX discrimination task (Chapter 4) was expected to tap into a higher level of phonological processing, since it includes a segmental match component in addition to a discrimination component. For segmental match, segments need to be recognized. It was expected that orthography could influence auditory processing at this higher level of phonological processing when segment recognition was required (Morais, Kolinsky, & Claes, 1995). The results

[41] Considering the effects of the factors 'vowel category' and 'lexicality' on the accuracy of discrimination, both factors [vowel category: $X^2(6) = 37.8$, $p < 0.0001$; lexicality: $X^2(4) = 10.8$, $p = 0.03$] and the interaction between them (p=0.02) showed significant values in the German discrimination task performed by German literates.

of the Persian auditory ABX discrimination task showed that minimal pairs with short vowels were discriminated as accurately as minimal pairs with long vowels by Persian literates. Thus, the mismatch between phonology and orthography in the case of vowels in Persian did not influence the discrimination accuracy in the ABX discrimination task. Moreover, minimal pairs whose items differed due to a vowel insertion (/?æmd-?æmud/) were discriminated less accurately than minimal pairs whose items differed due to a vowel replacement (/?æmud-?æmid/). According to Hahn and Bailey (2005), minimal pairs with phoneme insertion sound integrally more similar than minimal pairs with phoneme replacement, suggesting that participants performed the task at a perceptual level rather than at a recognition level. Any discrimination based on the recognition of spoken items and the segments included in them would lead to an almost perfect discrimination for minimal pairs differing in a vowel insertion. Also, minimal pairs differing in vowels with minimal perceptual distance (/xod-xud/) were discriminated less accurately than minimal pairs differing in perceptually less similar vowels (/bɑd-bæd/). This could also be an indication that the task was performed at a perceptual level rather than at a recognition level. Thus, both AX and ABX discrimination tasks were performed at a perceptual level by Persian literates, and (pseudo)words with short vowels were perceived as accurately as (pseudo)words with long vowels. Hence, in the case of vowels, orthographic features did not influence speech perception in Persian. These findings are at variance with interactive models for spoken and written word processing (McClelland, Mirman, & Holt, 2006; Grainger, Muneaux, Farioli, & Ziegler, 2005) in which it is assumed that speech perception can be shaped by orthographic knowledge through sublexical and lexical links between phonology and orthography or through top-down flow of information (Ziegler & Ferrand, 1998).

Further experiments in the Chapters three and five deal with the question whether orthographic knowledge can influence metaphonological processing of short and long vowels in the auditory modality. Persian literates who discriminated spoken stimuli with short vowels as accurately as spoken stimuli with long vowels were less accurate in manipulating segments of spoken words including short vowels as compared to segments of spoken words including long vowels in a Persian phoneme reversal task (Chapter 3). This finding suggests an influence of orthography on speech processing when explicit phonological processing is tapped into. Persian long vowels with orthographic representation seem to be retained better in memory for any further phonological operation (manipulation) compared to Persian short vowels, suggesting that phonological awareness to short vowels might be reduced in Persian literates due to their long exposure to Persian conventional writing.

German vowels with different qualities were manipulated regardless of the vowel category by German literates in the German phoneme reversal task (Chapter 3). German vowels with different qualities are consistently spelled, and the accuracy of manipulation was the same across vowel categories in this experiment, confirming that a phonological manipulation can be performed well on speech segments that are orthographically represented.

In a Persian phoneme monitoring task (Chapter 5), orthographically unrepresented short vowels were detected less accurately than orthographically represented long vowels by Persian literates. This result was limited to the short vowels with mid or short durations. The short vowel with long duration was detected as accurately as the long vowels, suggesting that when the input contains rich phonological information, detection can be performed based on this information alone. However, when the input contains less rich phonological information (in the case of short vowels with mid or short duration), detection seems to be performed considering both phonological and orthographic information. Thus, in the latter case, the decision whether the auditory input contained the target phoneme seems to take into account all available sources of information.

Persian illiterates were less accurate than Persian literates in the Persian AX discrimination task (Chapter 3) and failed to perform the ABX discrimination (Chapter 4), phoneme reversal (Chapter 3), and phoneme monitoring (Chapter 5) tasks. Literacy or exposure to written materials, thus, seems to facilitate speech discrimination, speech segmentation, and segment detection or manipulation. These abilities could not be sufficiently attained by the short training sessions before each experiment. Thus, as claimed in the literature reviewed in Chapter 2, phonological processing is different in literates and illiterates (Petersson, Reis, & Ingvar, 2001; Reis & Castro-Caldas, 1997).

The results of the reading experiments (Chapter 6) showed that written pseudowords with pointed short vowels were named less accurately than written pseudowords with long vowels. Also, vowels in later positions within pseudowords were more susceptible to be misnamed, and this effect was more pronounced in pseudowords with pointed short vowels. Considering that pseudowords need to be named by phonological recoding, this result suggests that memory span and phonological awareness to short vowels are reduced for short vowels as a result of long exposure to the Persian conventional writing. On the other hand, written words with short vowels were recognized as accurately as written words with long vowels in a semantic categorization experiment with written words (Chapter 6). Thus, whereas word naming has been reported to be affected by spelling transparency (e.g., (Rahbari, 2019)), our result suggests that semantic categorization is not. This differential effect of spelling transparency in naming and semantic categorization suggests that spelling transparency has an influence in tasks with phonological output, possibly due to 'orthographic cohorts' providing inconsistent phonological information (Seidenberg, Waters, & Barnes, 1984) with respect to short vowels. By contrast, in reading tasks with no phonological output, the performance may be based on other criteria (e.g., visual criteria), and not necessarily based on phonological criteria. Thus, the quality of visual processing is the same for words with short and long vowels, but the quality of phonological processing is different for these words.

It has been postulated that phonological representations can be restructured in terms of orthographic features (Taft, 2011). Hence, we expected potential orthographic effects on speech processing to be larger in participants who were more proficient in reading conventional writing. Yet, the performance in the two reading tasks (naming pseudowords and semantic categorization of real words) did not correlate with the performance in the Persian auditory ABX discrimination task (Chapter 6). This result could be an indication that different cognitive mechanisms may underlie reading and auditory speech discrimination. The evidence in the ABX discrimination task showed that the task was approached with an integral match strategy at a pre-attentive, perceptual level (see Figure 6). Whereas, reading is performed at an attentive cognitive level.

With respect to the inconsistent relationship between phonology and orthography in the case of Persian consonants, the results of an auditory primed lexical decision task (Chapter 7) showed that when corresponding primes and targets shared the same initial sound and spelling (consistent targets), lexical decision for such targets was faster than when the same targets were primed by unrelated word forms. By contrast, when corresponding primes and targets shared the same initial sound with different spellings (inconsistent targets), the lexical decision latency for such targets was equal to when the same targets were primed by unrelated word forms. Hence, there was only facilitatory phonological priming when primes and targets shared the spelling. This finding is evidence for an influence of orthographic knowledge on speech processing (consistency effect) at the lexical level in Persian. The consistency effect was not observed whenever an inconsistent target was encountered, but only for inconsistent targets in the context of phonological priming. Thus, phonological priming pre-activated consistent targets, but inconsistent targets only to a lower or no extent.

8-2. General discussion

In this part, I will discuss our findings on written and spoken word processing with respect to the current debates in each field and the prevalent models of written and spoken word processing.

8-2-1. Written word processing

To understand how orthographic information can influence spoken word processing one needs to take into account the reading process. Due to the limitations mentioned, the author of this dissertation could only administer two reading experiments, visual pseudoword naming and semantic categorization of written words, both on Persian. For the discussion of the reading process in Persian, therefore, the findings of experiments on word naming conducted by other authors (Rahbari, 2019; Bakhtiar & Weekes, 2015) will be taken into account as well.

8-2-1-1. Relating findings to current models of reading

In the domain of written word processing, the Dual Route Cascaded (DRC) model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) and the Bimodal Interactive

Activation model, BIAM, (Grainger, Muneaux, Farioli, & Ziegler, 2005) are considered to be the most relevant ones. In the DRC model, three routes are available when processing written input words. The non-lexical route works based on language-specific GPC rules which translate graphemes to corresponding phonemes. The resulting phonemes are shared with the phonemic level for production. In the lexical-semantic route, written input words are processed and mapped to lexical orthographic entries which activate the meaning of written input words. Activated entries in the semantic module can activate corresponding lexical phonological representations. In the lexical route, written input words are processed visually and mapped to lexical orthographic representations that can ultimately activate corresponding lexical phonological entries. The activated lexical phonological codes in the lexical and lexical-semantic routes are shared with the phoneme module for production. The lexical route in the DRC model allows for interactive activations which is also a feature of the BIAM. Because of this overlap, the findings on written word processing are discussed considering the DRC model only. Moreover, the original version of the BIAM (Grainger, Muneaux, Farioli, & Ziegler, 2005) has no production module. Since the performance on word and pseudoword naming experiments is to be interpreted, the findings of written word processing are discussed considering the DRC model that does have a production module.

8-2-1-2. How does the DRC model account for the findings on naming written words?

In experiments of reading aloud, Persian opaque words take longer to name than Persian transparent words (Bakhtiar & Weekes, 2015). Moreover, opaque words are less accurate than transparent words in naming (Rahbari, 2019). Together these findings constitute a spelling transparency effect in Persian. The contrast between opaque and transparent words is comparable to the results of experiments with irregular and regular words in other languages such as English, where irregular words, such as ‘pint’ /paint/, are named less accurately and more slowly than regular words such as ‘punt’. Irregularity is defined as a violation of the rules for grapheme-to-phoneme correspondences (GPC) reflecting the most frequent pronunciations of letters. For example, the dominant pronunciation for ‘i’ is /ɪ/; the pronunciation of ‘pint’ as /paint/ is, thus, considered to be irregular because it violates this GPC rule. Moreover, most irregular words are also inconsistent in that many words with similar spellings to those of irregular words are pronounced differently. Considering an irregular word such as 'pint', there are many other words with the same body unit ("grouping of letters in a monosyllabic word that comprises the terminal consonants plus the vowel" (Taft, 1992), p. 1004) which are pronounced differently such as 'hint- print- mint, etc.'

The Dual Route Cascaded (DRC) model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) explains the regularity effect by a conflict between the outputs of the lexical and the non-lexical routes that operate in parallel, and deliver their outputs to the phoneme level (see Figure 4, Section 1-3). The non-lexical route functions based on GPC rules from left to right. In the lexical route, lexical phonological codes are activated by corresponding lexical orthographic codes. The phoneme level receives contradictory phonological information for irregular words from these two routes. That is, for irregular words, incorrect phonological codes from the non-lexical route (e.g., /-ɪnt/ in 'pint') compete with the correct phonological codes from the lexical route (e.g., /-aint/ in 'pint'). Thus, when several phonemes (here /ɪ/ and /ai/) are active at the phoneme level, they inhibit each other leading to incorrect or slower replies in naming irregular words as compared to regular words for which the phoneme unit receives consistent phonological codes from the lexical and the non-lexical routes.

The DRC model explains the consistency effect in a different way. Information is assumed to be activated in cascades at various levels (features, graphemes, sub-word units (e.g., bodies), lexical orthographic representations, and lexical phonological representations (Taft, 1992)) in the lexical route. That means when a piece of information is activated at one level, it is spread to subsequent levels. Hence, as soon as body units of inconsistent words become active, they activate corresponding phonological codes. For example, the body '-ood' activates /-u:d/ and /ʊd/ in words such as 'mood, good, wood, etc.' These activated phonological codes provide inconsistent pronunciations leading to a competition which results in slower and less accurate naming responses for inconsistent words. For consistent words, however, a consistent pronunciation is activated by the body unit. Thus, the regularity and consistency effects are the results of between-routes and within-route competition respectively.

Coming back to Persian, the outputs of the non-lexical route are underspecified as compared to the outputs of the lexical route with respect to the vowels of opaque words. The same is not the case for transparent words. For opaque words, the non-lexical route provides the pronunciation of consonants only, whereas the lexical route provides complete pronunciations including consonants and vowels. For transparent words, the non-lexical and the lexical routes provide pronunciations with specified vowels and consonants. The observed spelling transparency effect in Persian cannot be interpreted as being due to some kind of irregularity as in opaque words, there are no graphemes for short vowels to be converted to corresponding phonemes in the non-lexical route; thus, the GPC rules of Persian are not violated. Nonetheless, due to the lack of short vowel graphemes, Persian readers cannot rely on the non-lexical route to name opaque words, and opaque words can only be named via the lexical route. Considering that only 11% of Persian written words are completely transparent and the rest are either partially or totally opaque (Bakhtiar, Mokhlesin, Pattamadilok, Politzer-Ahles, & Zhang, 2021), it can be assumed that, in practice, only two routes are available to Persian readers, the lexical and the lexical-semantic routes (Bakhtiar & Weekes, 2015).

Within the lexical route, transparent words have advantages over opaque words. The first advantage is related to vowels embedded in bodies, which are specified in transparent words and underspecified in opaque words. Considering opaque words with graphemes that represent consonants only (e.g., <srm> /serom/ ‘serum’), it is not clear what the vowels of bodies are. In contrast, in the bodies in transparent words (e.g., <bad> /bɑd/ ‘wind’), the vowels are specified. In other words, the DRC model cannot explain on what basis activation spreads between units in the lexical route for opaque words, if bodies are defined as ‘the letters representing ultimate consonants and the vowel (Taft, 1992)’. One might assume that in opaque orthographies such as Persian, bodies are simply the ultimate letters or graphemes. Written inputs might then activate lexical entries sharing ultimate graphemes, i.e., bodies (e.g., <-rm> or <–ad>) in addition to lexical entries sharing the onsets of written inputs (i.e., initial grapheme <s-> and <b->).

The second advantage of transparent words over opaque words in the lexical route is related to the consistency of activated phonological codes following the activation of bodies for opaque and transparent words. Since bodies are not specified with respect to the embedded vowel in opaque words, the corresponding activated phonological codes are inconsistent for opaque words as compared to those for transparent words which have specified vowels. For example, the body <–rm> which is silent about the embedded vowel, activates many phonological codes in the lexical route, among others /-orm, -erm, -erem, -erom, -oræm, -ærm/ in the words '<jrm> /ʤorm/ 'crime', <jrm> /ʤerm/ 'mass', <krm> /kerm/ 'worm', <krm> /kerem/ 'creme', <srm> /serom/ 'serum', <khrm> /xoræm/ 'fresh', <nrm> /nærm/ 'soft'. Consequently, phonological codes are inconsistent with respect to the vowel(s) included. In the case of transparent words, the activated phonological codes following the activation of bodies are specified and consistent with respect to body vowels. For example, <-ad> will activate only /–ɑd/ in words such as '<bad> /bɑd / 'wind', <had> /hɑd / 'acute', <dad> /dɑd / 'scream', etc. Hence, the activated phonological codes for bodies are consistent in transparent words. Thus, the issue with opaque words is twofold: the underspecified vowels and the competition among a large number of inconsistent phonological codes for bodies, both lead to a less accurate and slower naming of opaque words as compared to transparent words.

The phonological inconsistencies for opaque words might be resolved later by the activated onset nodes (Taft, 1992). For example, let's assume that the written input <khrm> activates all lexical orthographic entries having <kh-> in the onset and <-rm> in the body. The body <-rm> activates these phonological codes /-orm, -erm, -erem, -erom, -oræm, -ærm/. The opaque word /xoræm/ would then be activated more strongly than the competing words containing the other possible phonological codes of the body <-rm> due to the additional activation it receives from the activated onset nodes <kh-> /x-/. In this way, the size of the set of competing opaque words can be reduced via the activation they receive from the activated onset nodes.

Thus, although the DRC model can account for regularity and consistency effects in naming in other languages, the assumed mechanisms for these effects cannot satisfactorily account for the spelling transparency effect in Persian. Rather than being due to incongruent or inconsistent phonological codes generated in the sub-lexical route via GPC rules, spelling transparency effects in Persian are better explained as arising in the lexical route due to underspecified bodies in opaque words with respect to short vowels.

8-2-1-3. How does the DRC model account for different effects of spelling transparency across tasks?

As mentioned earlier, the spelling transparency effect has been mainly observed in tasks with phonological outputs e.g., naming written words. In tasks without phonological outputs e.g., lexical decision or semantic categorization, the spelling transparency effect has either not been observed, or only observed for low frequency words in some studies (Baluch, 1993; Koriat, 1985). As the spelling transparency effect emerges as a result of the competition among inconsistent phonological codes, it can be explained by the DRC model why such effects are absent in tasks without phonological output. In semantic categorization tasks, responses can be made via the lexical-semantic route based on visual and semantic criteria. In lexical decision tasks, responses can be made based on visual criteria via the lexical route. Thus, phonological codes are only necessary in naming tasks. Since all these tasks are normally run under time limitation in experimental contexts, the recourse to unnecessary processes or criteria may be avoided. As soon as replies are available, processing may be terminated.

With respect to the effect of spelling transparency for low frequency opaque words in tasks without phonological outputs, one may assume that the visual identification of low frequency words is slower, so that inconsistent phonological codes for subcomponents of opaque words (i.e., bodies) may become available before they can be identified or recognized visually (Seidenberg, Waters, & Barnes, 1984).

8-2-1-4. What does the spelling transparency effect show?

As argued above, the spelling transparency effect in Persian seems to arise in the lexical route. The difference in performance between opaque and transparent words may, thus, be informative about the organization of lexical entries or the way the activation spreads through lexical entries. Written input words activate all lexical orthographic entries sharing their onset or body. Thus, lexical orthographic entries with the same body unit may be interconnected. Such connections may be formed based on the similarity in spelling and sound. This means that lexical items with the same body unit may form different sub-networks if the body unit corresponds to several phonological codes (Siew & Vitevitch, 2019). For example, written words with the body unit –ood may enter into different sub-networks based on the pronunciations they provide. In these networks, ‘good, wood, etc.’ are in the same sub-network and ‘mood’ in another sub-network. Those in the first network may have a higher resting activation level and be named faster due to having the more frequent pronunciation as compared to those in the second network (Cortese & Simpson, 2000). Thus, the inconsistency in pronunciation for similarly spelled units might organize the structure of the lexical orthographic entries. Moreover, the structure of lexical networks may influence the way they are processed (Vitevitch, Goldstein, Siew, & Castro, 2014). The lexical activation can be performed based on networks rather than based on individual lexical entries. Coming across a written input word with the body unit –ood, the network including lexical items with the same body unit as the input and including the most frequent pronunciation (e.g., good, wood, etc.) may be activated first or more strongly while processing written words. Only afterwards or more weakly, the network that includes lexical items with the same body unit as the input, but the less frequent pronunciation (e.g., mood) may be activated. Thus, lexical entries with similar spelling and sound may be stored more closely than lexical entries with similar spelling but dissimilar sounds, and lexical activation may be performed with a priority for the network with denser lexical items.

8-2-1-5. How does the DRC model account for the findings on pseudoword naming?

In the DRC model, non-lexical route processing can be assessed by pseudoword naming, because pseudowords can only be pronounced via the non-lexical route. The findings of our pseudoword naming experiment with long vowels or pointed short vowels showed (i) an equal performance for short pseudowords with one long vowel or one pointed short vowel (ii) a less accurate performance for pseudowords with three pointed short vowels compared to pseudowords with three long vowels, and (iii) less accurate naming of vowels in later positions than vowels in earlier positions. The latter effect was more pronounced for pseudowords with pointed short vowels compared to pseudowords with long vowels.

According to the DRC model, deficits in successful reading may have proximal or distal causes (Jackson & Coltheart, 2001). Proximal causes refer to those which reside in the reading system itself. Distal causes reside outside the reading system (e.g., in language or cognitive processing), but they can influence the successful functioning of the reading system. One of the distal causes leading to unsuccessful reading is related to phonemic awareness (Jackson & Coltheart, 2001), i.e., awareness of phonemes in spoken words (Castle, 2006). Phonemic awareness can particularly influence the function of the non-lexical route, because some ability in segmenting and assembling phonemes is required for the successful performance in this route (González-Valenzuela, Díaz-Giráldez, & López-Montiel, 2016).

The results of the pseudoword naming experiment showed that the problem in naming long pseudowords with pointed short vowels cannot be related to a proximal cause residing in the non-lexical route. If Persian readers had problems in converting short vowel diacritics to the corresponding phonemes (i.e., problem in performing GPC), they should have named short pseudowords with one pointed short vowel less accurately than short pseudowords with one long vowel, which was not the case. Thus, the category of the vowel in pseudowords and the length of pseudowords both contributed to the unsuccessful performance in naming suggesting that short vowels cannot be retained long enough in memory to utter due to lower phonemic awareness for these vowels. Thus, a distal cause, i.e., phonemic awareness, seems to be the most plausible reason for the observed performance in the pseudoword naming experiment.

8-2-2. Spoken word processing

Orthographic effects on spoken word processing refer to the observation that spoken words with consistently spelled phonological units are processed faster or more accurately than (i) spoken words with inconsistently spelled phonological units, (ii) spoken words in which different phonological units are spelled the same, and (iii) spoken words whose phonological units have no orthographic realization (Stone, Vanhoy, & Van Orden, 1997). In this section, I will discuss the findings of this dissertation on orthographic effects in Persian with respect to current models of spoken word processing: I will, furthermore, consider the current debates on the nature of orthographic effects and the level(s) at which these effects are observed. Moreover, I will elaborate on how investigating orthographic effects can shed more light on the interaction debate in the domain of spoken word processing.

8-2-2-1. Relating findings to models of spoken word processing

Among the models of spoken word processing, two models consider the role of orthographic knowledge on spoken word processing, the BIAM (Grainger, Muneaux, Farioli, & Ziegler, 2005) and the Merge model (Norris, 1999).

According to the BIAM, processing spoken words coactivates orthographic information in addition to phonological information at the prelexical and lexical levels. Therefore, orthographic effects are expected to be online and mandatory at those levels. According to the Merge model, orthographic effects are expected to be observed at the post-lexical level whenever orthographic information facilitates spoken word processing. Hence, the role of orthographic information in spoken word processing is assumed to be complementary rather than mandatory in the Merge model. In this model, the orthographic effect is not an online (real-time) effect, but it emerges when the decision unit at the post-lexical level reconsiders the information from auditory inputs with respect to other sources of knowledge e.g., orthographic knowledge. Thus, an orthographic effect is not an effect which occurs in normal speech processing. Rather, it emerges in experimental contexts, or when spoken input words cannot be identified by their phonological information alone.

Other models of spoken word processing e.g., the Cohort model (Marslen-Wilson & Tyler, 1980) or the TRACE model (McClelland & Elman, 1986), do not consider any role for orthographic knowledge in spoken word processing.

8-2-2-1-1. Interpreting the findings based on the Cohort and TRACE models

We found orthographic effects in the Persian auditory primed lexical decision experiment and the German AX discrimination experiment. Spelling inconsistency between phonologically related primes and targets modulated the phonological priming effect found for phonologically related primes and targets with consistent spelling in the Persian primed lexical decision experiment performed by Persian literates. In the auditory German AX discrimination experiment, German homographic word pairs were discriminated less accurately than German non-homographic word pairs when the word pairs contained the same vowel contrast. Hence, orthographic information, in addition to phonological information, had an influence on spoken word processing. These findings are not consistent with the assumption of the Cohort and TRACE models that spoken inputs are processed based on their phonological information alone.

8-2-2-1-2. Interpreting the findings based on the Bimodal Interactive Activation Model (BIAM)

On the other hand, the findings of this dissertation suggest that orthographic information is not co-activated mandatorily while spoken input words are processed, hence, the findings are also not consistent with the BIAM. If orthographic information was mandatorily activated in spoken word processing, words with inconsistently spelled phonological units would be expected to be processed more slowly than words with consistently spelled phonological units in any context, due to the online activation of multiple orthographic codes for words with inconsistently spelled phonological units and the inhibitory competition among them. However, in the primed lexical decision task, no orthographic effect was observed when words with consistently or inconsistently spelled phonological units were preceded by phonologically unrelated word forms. In the same task, the orthographic effect was observed when spoken words with consistently or inconsistently spelled phonological units were preceded by phonologically related word forms. Therefore, spelling consistency or inconsistency as such does not seem to have triggered the orthographic effect, rather this effect was only triggered in the context of phonological priming.

Moreover, the orthographic effect was observed for words, but not for pseudowords in the German AX discrimination experiment suggesting that the orthographic effect is limited to the lexical level. The BIAM cannot explain this pattern. According to this model, the orthographic effect should have been observed for inconsistent words as well as for inconsistent pseudowords due to the sub-lexical similarity between words and pseudowords and bilateral links between phonological and orthographic representations at the sub-lexical level.

Furthermore, according to the BIAM, the perception of spoken words is assumed to be shaped by orthographic information through a top-down flow of information (Ziegler & Ferrand, 1998). The evidence from our Persian AX and ABX discrimination experiments does not support this assumption. In these tasks, spoken (pseudo)words containing phonological units with or without orthographic realization were discriminated equally accurately suggesting that the orthographic contrast did not influence spoken (pseudo)word discrimination. The evidence in these experiments showed that stimuli were discriminated at the perceptual level: (Pseudo)words differing in an inserted vowel (e.g., /kerem-kerm/) were discriminated less accurately than (pseudo)words differing in vowel replacement (e.g., /ʤerm-ʤorm/) suggesting that dissimilarities between (pseudo)word-pairs were only attended when they were perceptually more salient, i.e., in pairs differing with respect to vowel replacement (Hahn & Bailey, 2005).

Thus, orthography influenced word discrimination when identification-recognition knowledge, i.e., lexical knowledge (German AX discrimination), rather than perceptual knowledge (Persian AX and ABX discrimination) was used to discriminate suggesting that the orthographic effects had a lexical origin. Discriminating words at different processing levels in German and Persian can be interpreted considering the vowel system density in these two languages. That is, in languages with a dense vowel system such as German, when words differing in vowels cannot be reliably differentiated perceptually, lexical knowledge may come into play (Escudero, Hayes-Harb, & Mitterer, 2008). At the lexical level, words with dissimilar sounds but similar spellings may be stored more closely than words with dissimilar sounds and spellings leading to the emergence of an orthographic effect in German, but not in Persian with its sparse vowel inventory that can be reliably differentiated perceptually. Finding an orthographic effect at the lexical, but not at the prelexical or perceptual level is at variance with the BIAM which assumes that lexical knowledge can flow top-down and that phonological and orthographic representations interact at the lexical as well as at the pre-lexical level.

8-2-2-1-3. Interpreting the findings based on the Merge model

According to the Merge model, orthographic information is intentionally recruited in addition to phonological information when orthographic information can facilitate the performance in auditory tasks. We observed orthographic effects in the German AX discrimination and the Persian primed lexical decision experiments. In both experiments, orthographic information led to poorer performance. In the German AX discrimination experiment, German listeners discriminated spoken words less successfully when they were homographic. In the Persian primed lexical decision experiment, a phonological priming effect observed for corresponding primes and targets with similar sound and spelling was neutralized for corresponding primes and target with similar sound but dissimilar spelling. These findings are difficult to reconcile with the Merge model's assumption of a strategic recruitment of orthographic knowledge in phonological processing for a number of reasons. First, if a piece of information is used strategically or intentionally to improve performance, it can also be controlled or avoided intentionally when it impairs performance. The performance of participants in the German AX discrimination and Persian primed lexical decision experiments showed that the effect of orthography could not be intentionally avoided in these auditory tasks when it impaired performance. Second, orthographic knowledge can be used in difficult auditory tasks (e.g., rhyme detection) to improve performance due to the benefit of orthography in retaining auditory information in memory (Castle, Holmes, Neath, & Kinoshita, 2003). Spoken word discrimination, however, is one of the easiest paradigms in auditory processing which does not load memory or does not require lexical access (Gerrits & Schouten, 2004), and hence is unlikely to involve strategic factors. Nonetheless, an orthographic effect was observed in the German AX discrimination task. Third, intentional activation of other sources of knowledge in a tight experimental timing when they are not necessary or facilitatory for task performance does not seem plausible. Fourth, according to the Merge model, Orthographic knowledge can be used in auditory word processing when orthographic knowledge is rendered salient. In the German AX discrimination task, just 6% of the whole stimuli set was composed of homograph words. Such a low proportion cannot render the orthographic knowledge salient. Taken together, our evidence suggests that the recruitment of orthographic knowledge leading to orthographic effects was 'unavoidable' rather than strategic in these tasks (more arguments are provided in the Section 8-2-2-2).

8-2-2-1-4. The nature of lexical associations

Almost all models of spoken word processing, e.g., Cohort, TRACE, Merge, and BIAM, maintain that spoken word processing is incremental. Listeners do not wait until the end of spoken words to start processing them, rather, several lexical entries consistent with the segments of spoken inputs are activated (Weber & Scharenborg, 2012). This notion of parallel activation suggests that ‘cap’ activates ‘capital’ and ‘captain’, for example. Also ‘captain’ and ‘capital’ activate the embedded word ‘cap’. Moreover, ‘care’ can activate ‘cat’ and ‘bare’. The mental lexicon is, hence, structured in a way that there are similarity associations among lexical entries, and similar entries are stored more closely leading to quicker recognition of e.g., embedded words (e.g., ‘cap’) when presented after their embedding words (e.g., ‘captain’ and ‘capital’).

Considering the notion of parallel activation and similarity associations between lexical entries, the findings of this dissertation are consistent with the assumptions of current spoken word processing models. However, the findings of this dissertation cast doubt on the assumptions of these models with respect to how similarity association is defined. Whereas the models of spoken word processing assume phonological, i.e., sound-based, associations among lexical phonological entries (Weber & Scharenborg, 2012), the findings of this dissertation suggest that lexical phonological entries are associated phonologically and orthographically (phonographic associations). The results of the primed lexical decision experiment indicated that the residual activation of embedded target words following the recognition of embedding primes was stronger when embedded and embedding words shared similar spellings compared to when embedded and embedding words shared dissimilar spellings suggesting phonographic associations among lexical entries rather than mere phonological associations. If lexical phonological entries were merely phonologically associated, orthographic associations between embedded and embedding words would play no role, and the residual activation of embedded words following the recognition of embedding words with similar or dissimilar spelling would have been equal. Moreover, in the German AX discrimination task, discriminating words with dissimilar sounds but similar spellings were less accurate than discriminating words with dissimilar sounds and spellings suggesting that phonographic associations among words matter.

8-2-2-2. What is the nature of orthographic effects?

As explained earlier in Chapters 1 and 2, there is no consensus on the nature of orthographic effects in spoken word processing. A group of authors claim that orthographic information is mandatorily activated at the lexical and prelexical levels in real time while processing spoken words, i.e., ‘*online* account’ (Pattamadilok, Perre, Dufau, & Ziegler, 2009; Ziegler, Petrova, & Ferrand, 2008; Taft, Castles, Davis, Lazendic, & Ngye-Hoan, 2008; Perre & Ziegler, 2007). Other authors claim that phonological information does not activate orthographic information while processing spoken words but, instead, phonological representations are restructured in terms of orthographic features, i.e., '*offline* account' (Taft, 2011; Perre, Pattamadilok, Montant, & Ziegler, 2009; Taft & Hambly, 1985). For example, the rhyme in 'swap' may develop two phonological representations /-ɔp/ and /-æp/ after learning to read and write (Taft, 2011). A third group of authors claim that the influence of orthography on spoken word processing is strategic or intentional to facilitate the performance in auditory experiments (Cutler & Davis, 2012; Cutler, Treiman, & Van Ooijen, 2010). A fourth group of authors maintains that orthographic information may restructure the organization of lexical phonological entries in such a way that lexical entries with similar sounds and spelling reside in the same network, but lexical entries with similar sound but dissimilar spelling reside in different networks, i.e., '*network* account' (Siew & Vitevitch, 2019; Muneaux & Ziegler, 2004). According to this account, processing words residing in the same network is faster and more accurate than processing words residing in different networks. Thus, the '*network* account' functions based on the phonographic similarity associations between lexical entries. Findings of an influence of orthography on spoken word processing have been mainly discussed based on the first three accounts. The '*network* account' has so far received less attention in this respect, possibly because it has emerged more recently drawing from sociology, computer science, and a number of other fields to understand the structure of mental lexicon (Vitevitch, Goldstein, Siew, & Castro, 2014). According to the '*network* account', orthographic effects are structural-residual effects of an offline nature (Muneaux & Ziegler, 2004). That is, learning to read and write can restructure the organization of lexical phonological entries in the mental lexicon, but not the nature of the lexical phonological representations as such, as assumed by the '*offline* account'. I will now discuss how compatible the findings of this dissertation are with these accounts.

In the primed lexical decision task, consistent target words were processed faster when they occurred after words with similar sound and spelling (e.g., /keʃvær - **keʃ**/, کِشور - **کِش**) than when the same targets occurred after unrelated word forms (lægæn-**keʃ**). By contrast, inconsistent target words were not processed faster when they occurred after words with similar sound but dissimilar spelling (e.g., /tænɑb - **tæn**/, **تِن**-طِناب) compared to when the same targets occurred after unrelated word forms (kuʧeh-**tæn**). At the same time, inconsistent targets were processed as fast as consistent targets after unrelated primes. This pattern of results is at variance with the online, *offline*, and *strategic* accounts. In the case of online or offline orthographic effects, inconsistent targets should have been processed more slowly than consistent targets after unrelated as well as related word forms due to the competition between different spellings for the same inconsistently spelled phonological unit, or due to developing multiple underspecified phonological representations for an inconsistently spelled phonological unit, which was not the case. Furthermore, the performance in the primed lexical decision experiment suggests that the orthographic effect was not due to the strategic recruitment of orthographic knowledge. Strategies are supposed to provide additional cues for making responses, or to facilitate the use of working memory in auditory tasks (Taft, Castles, Davis, Lazendic, & Ngye-Hoan, 2008). Accordingly, if an orthographic strategy (i.e., activating orthographic information in addition to phonological information) was adopted in the primed lexical decision experiment, this strategy should have remained engaged throughout the whole experiment, i.e., in contexts with and without phonological priming. By contrast, the evidence shows the orthographic effect only with phonological priming. Moreover, the orthographic effect led to slower rather than faster reaction times in conditions with phonological priming. As discussed above, this result makes the employment of a strategy less likely as strategies are employed to improve or facilitate performance.

In the German AX discrimination experiment, homographic minimal pairs differing in vowel length were discriminated less accurately than non-homographic minimal pairs differing in vowel length. This difference was not observed for homographic and non-homographic pseudoword minimal pairs. An orthographic effect as assumed by the *online* and *offline* accounts should have been triggered for both words and pseudowords. For example, the units /-ɑxən/ and /-ɑ:xan/ <-achen> should have been activated when listening to the words ‘lachen - Lachen’ as well as when listening to the pseudowords /pɑxən / and /pɑ:xən/ <-achen> (Taft, 2011). Restructured phonological representations should have led to impaired discrimination for potentially homograph pseudowords due to the top-down flow of information, which was not the case. According to the *offline* account, either of the homographs 'lachen - Lachen' would develop two phonological representations that are not well specified with respect to the first vowel /-ɑ(ɑː)xən/. Such restructured and less specified phonological representations should have led to poorer discrimination of homographic pseudowords with the same vowel contrast due to a top-down flow of information. Furthermore, the sublexical links between phonology and orthography could impede the discrimination of homographic pseudoword minimal pairs, based on the *online* account. In sum, the *online* and the *offline* accounts do not seem to make the right predictions for the German AX discrimination task.

Also, the strategic account is not without problems when considering our results. This account assumes that orthographic knowledge is considered at a post lexical level in addition to phonological knowledge to facilitate performance in difficult auditory tasks. This process is assumed to be intentional, thus it could be avoided or controlled in case it leads to poorer performance. It may be argued that orthographic knowledge is intentionally activated as an aide to phonological knowledge in transparent orthographies because orthography reflects phonology in most cases. This assumption would, however, not explain our results of the German AX discrimination experiment, because if a source (i.e., orthography) is used as an aide to phonological knowledge, that source should be used always, that is irrespective of whether phonological units occur in words or in pseudowords. Arguably, the additional knowledge source might even be expected to be used for unfamiliar phonological units (i.e., in pseudowords) more frequently than for familiar phonological units (i.e., in words) to improve performance for the more demanding unfamiliar units (Escudero, Hayes-Harb, & Mitterer, 2008).

In sum, the findings of this dissertation are not fully compatible with *online*, *offline*, or *strategic* accounts, but seem to be more in line with the '*network* account'. According to this account, lexical phonological entries that are phonologically associated are more strongly associated when they are spelled similarly than when they are not. In the primed lexical decision task, embedded words were recognized faster after the recognition of embedding words when they shared spellings compared to when they had different spellings suggesting that orthographic features can restructure lexical phonological entries to be more closely related in the former case. According to the '*network* account', cognitive processes that are performed on a network depend on the structure of the network (Siew & Vitevitch, 2019; Vitevitch, Goldstein, Siew, & Castro, 2014). Lexical entries spread their activation to entries in the close neighborhood more strongly than to entries in the far neighborhood. Moreover, the '*network* account' can accommodate why the orthographic effect was observed with phonological priming but not without phonological priming in this task, because of the structural residual effects which are the case in the condition with phonological priming but not in the condition without phonological priming. That means, the orthographic effects, following the '*network* account', emerge when words are presented in community, or else when the experimental paradigm can reveal those processing stage(s) involved in the spread of activation among lexical entries or competition among lexical entries for selection.

Also, according to the '*network* account', words with dissimilar sounds but similar spellings stay more closely than words with dissimilar sounds and spellings. The findings of the German AX discrimination task showed that homograph minimal pairs sounded more similar than the non-homograph minimal pairs with the same vowel contrasts suggesting that spelling similarity brings words with dissimilar sounds closer. By contrast, spelling dissimilarity pushes words with dissimilar sounds away in the lexical network. Also, the '*network* account' can explain why the orthographic effect was found for words but not for pseudowords in the German AX discrimination experiment. Since the '*network* account' functions based on the phonographic structure of lexical entries, pseudowords with no exact lexical entry are not expected to show a strong orthographic effect, though a mild orthographic effect cannot be refuted for word-like pseudowords when used in the community of phonologically similar words and/or when intentional segmentation of spoken pseudowords is induced.

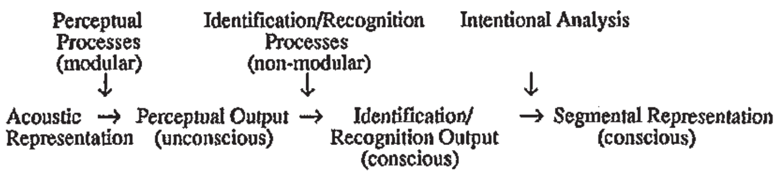
For example, Taft et al., (Taft, Castles, Davis, Lazendic, & Ngye-Hoan, 2008) found a facilitatory phonological priming effect when target words were preceded by pseudo-homograph primes (e.g., /driːd-dred/ <dread>) but the same facilitation was not observed when target words were preceded by non-pseudohomograph primes (e.g., /ʃriːd-ʃred/ <shread-shred>) in an auditory lexical decision task. Here, the pair with dissimilar sounds but similar spelling would activate lexical entries that are closer together based on the '*network* account', therefore, lexical decision is facilitated. The pair with dissimilar sounds and spellings would activate lexical entries that are farther according to the '*network* account', therefore, lexical decision is not facilitated. In most of the studies, the orthographic effects are not reported when pseudowords were used in trials of simple lexical decision, i.e., not in word community within a trial (e.g., (Ventura, Morais, & Kolinsky, 2007; Ventura, Morais, Pattamadilok, & Kolinsky, 2004; Ziegler & Ferrand, 1998)) or when pseudowords were not word-like (see (Taft, 2011) for this argumentation).

The findings of this dissertation are better interpretable based on the '*network* account' which assumes that orthographic effects are structural residual effects and emerge unavoidably through restructuring lexical entries in terms of phonographic features. But, it might be argued that awareness played a role in the selection of lexical candidates. During the recognition of embedding primes, participants might have deliberately focused attention on some activated lexical entries that were related to embedding primes in sound and spelling; hence, those entries with similar sounds but dissimilar spellings to those of corresponding primes were inhibited. This option can be plausible if it is assumed that lexical selection is a two-stage process (Simpson & Burgess, 1985) in which the activation is first automatically spread through all similar sounding lexical entries followed by an optional stage driven by attentional resources which inhibits similar sounding lexical entries with dissimilar spelling to that of spoken inputs. However, in this case, it was expected to find a spelling dominancy effect, which means the inhibition of lexical entries with subordinate spelling should have been easier than the inhibition of lexical entries with dominant spelling in the related trials, which was not the case. In the primed lexical decision experiment, when inconsistent targets contained subordinate or dominant spellings, lexical decision latencies did not differ in the related trials suggesting that the orthographic effects cannot be strategic or intentional.

8-2-2-3. Which levels of speech processing are orthographically influenced?

All models (Cohort, TRACE, and BIAM), except for the Merge model, are silent about metaphonological processing of spoken inputs, i.e., performing intentional operations on speech segments, and consequently also about possible orthographic effects in metaphonological processing. In the Merge model, the decision unit has been incorporated to account for intentional processing of speech in auditory experiments such as phoneme monitoring experiments. This model predicts orthographic effects in difficult auditory tasks with a metaphonological component. Our findings, however, suggest that orthographic effects are not limited to difficult tasks with a metaphonological component. We also observed orthographic effects in easy to perform tasks such as the German AX discrimination task.

Figure 6 (repeated below for the convenience of the reader) shows the speech processing levels assumed by Morais and colleagues (Morais, Kolinsky, & Claes, 1995). The authors assume that of the three processing levels (perceptual, identification/recognition (i.e., lexical), and intentional), the perceptual level is resistant to orthographic effects, whereas the others can be influenced by orthographic features. The findings of this dissertation are in agreement with this view.



Our results showed orthographic effects at the lexical (identification-recognition) level (German AX discrimination experiment and Persian primed lexical decision experiment) and at the intentional level (Persian phoneme reversal and phoneme monitoring experiments). At the perceptual level, the orthographic effect was not observed (the Persian AX and ABX discrimination experiments).

The findings of this dissertation have been discussed in previous sections with respect to the experiments targeting perceptual and lexical levels. As for an orthographic effect at the intentional level, our findings showed that phonological units without orthographic realization were manipulated less accurately than phonological units with orthographic realizations in the phoneme reversal experiment. In this experiment, we could not refute the adoption of a visual strategy while performing the task. Phonemes could have been orthographically visualized, and, in consequence, metaphonological processing could have been performed better for phonemes that have orthographic representations. Moreover, a visual strategy might plausibly be adopted in a difficult task such as phoneme reversal. Hence, in a separate experiment, we administered an easier task to investigate the effect of orthography at the intentional level. In this phoneme monitoring experiment, we found that, when vowels were of long duration, they were detected more accurately in words than in pseudowords irrespective of whether they had an orthographic realization or not. When vowels were of mid or short duration, they were detected more accurately when they were orthographically represented than when they were not. This result pattern suggests that participants adopted a visual strategy in conditions with weaker phonological cues. Vowels of mid or short duration could be detected by segmenting the constituents of spoken items based on graphemic information. This visual segmentation was more accurate for spoken items whose constituents were orthographically represented than for spoken items whose constituents were not. By contrast, vowels of long duration could be detected based on phonological cues alone, that is without visual segmentation. Our results suggest that at the intentional level, sounds can be better manipulated and detected when orthographically represented. Listeners are able to recruit orthographic information or not at the intentional level, depending on the quality of the phonological information and phonological processing.

It should be noted that the nature of the orthographic effects is different at the lexical and intentional levels. At the lexical level, the orthographic effects are structural residual effects and emerge unavoidably through restructuring lexical entries in terms of phonographic features, whereas at the intentional level, orthographic representations may be used strategically to aid memory in the intentional segmentation of spoken items.

8-2-2-4. Is the architecture underlying spoken word processing autonomous or interactive?

According to interactive models (Figure 23, a), lexical information can influence pre-lexical processes via a top-down flow of information (McClelland, Mirman, & Holt, 2006). In contrast, autonomous models (Figure 23, b) assume that information flows bottom-up only, and lexical and perceptual (pre-lexical) information is integrated at a decision level (Norris, 1999). Moreover, there is an interaction between phonological and orthographic information according to BIAM (Grainger, Muneaux, Farioli, &

Ziegler, 2005). Ziegler and colleague (Ziegler & Ferrand, 1998) found that feedforward (spelling-to-sound) and feedback (sound-to-spelling) inconsistencies affect word perception. The authors interpreted this finding as support for interactive models. Note, however, that if orthographic effects were due to a feedback mechanism, such effects should be observed for both words and pseudowords. Ziegler and colleague, however, found the orthographic effect (sound-to-spelling consistency) only for words but not for pseudowords in a lexical decision task. They attributed this result to different mechanisms used to decide whether spoken items are real words or pseudowords. When the activation level of a lexical entry passes a required threshold due to being the most similar entry to a spoken input, it is recognized as a real word, however when no lexical entry reaches the required activation threshold and the time is out, the input is recognized as a pseudoword. According to Ziegler and Ferrand (1998), this time-out mechanism can potentially conceal the orthographic consistency effects for pseudowords, because responses based on the time-out mechanism for pseudowords can be made before any consistency effects can emerge.

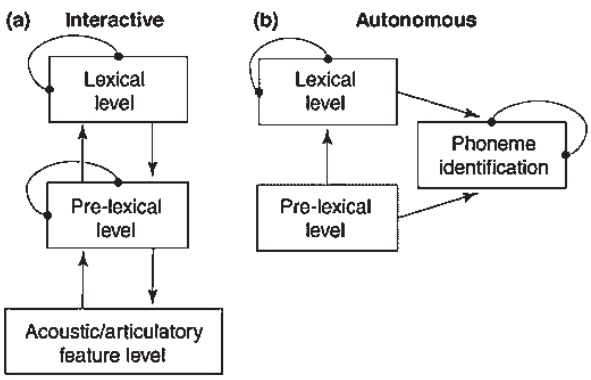


Figure 23. Schematic representation of information flow in interactive (a) and autonomous (b) models with arrows showing excitatory links and curves showing inhibitory links (McClelland, Mirman, & Holt, 2006), p.13.

Taft (2011) criticized Ziegler and Ferrand's explanation for the failure to find the orthographic effects for pseudowords. According to Taft, the time-out mechanism provides enough time for the feedback mechanism to trigger and the orthographic effects to emerge for words and pseudowords and, thus, cannot explain the failure to find an orthographic effects for pseudowords reported by some studies (e.g., (Ventura, Morais, & Kolinsky, 2007; Ventura, Morais, Pattamadilok, & Kolinsky, 2004; Ziegler & Ferrand, 1998)) and also found in our German AX discrimination experiment. The lack of orthographic effects for pseudowords, thus, remains a challenge for interactive models.

Moreover, if orthographic effects were the result of an interaction between lexical and pre-lexical phonological and orthographic representations, they should arise independently from context, in our case independently from phonological priming. However, in our primed lexical decision task, no orthographic effect was observed with unrelated primes, suggesting that an interaction mechanism cannot be the (sole) source of orthographic effects on spoken word processing.

These considerations suggest that even though at early stages of learning to read and write, there must be interactions between lexical and pre-lexical phonology and orthography leading to gradual restructuring of lexical entries in terms of phonographic features; after a long exposure to written language, interactions may dwindle in adult skilled readers. In proficient readers, restructured lexical phonographic knowledge may serve lexical operations with no feedback to the pre-lexical level. Thus, a flexible model which allows for interactions at early stages of reading and writing and the autonomy of systems at later stages can accommodate the findings.

8-2-2-5. Literacy and spoken word processing

Among the tasks dealing with (meta)phonological processing of spoken words, illiterates were only able to perform auditory AX discrimination, but none of the other tasks (auditory ABX discrimination, phoneme reversal, and phoneme monitoring). This performance of Persian illiterate participants confirms reports from other authors (Petersson, Reis, & Ingvar, 2001; Petersson, Reis, Askelöf, Castro-Caldas, & Ingvar, 2000) that at least in experimental conditions, phonological processing is qualitatively different between illiterate and literate participants, suggesting that explicit access to word segments requires learning to read and write (Morais, Cary, Alegria, & Bertelson, 1979).

Chapter 9: Conclusions and suggestions for further research

9-1. Conclusions

This dissertation investigated written and spoken word processing with the main focus on the latter. Our findings have shown that inconsistencies in phonological information due to the lack of orthographic realizations for Persian short vowels influences written word processing, while the inconsistencies in spelling for the same phonological unit influence spoken word processing. These findings have theoretical implications for current models of words processing.

Models of written word processing (e.g., DRC (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001)) consider the role of phonology in written word processing, but cannot account satisfactorily for all aspects of our findings, i.e., how activation spreads among entries in orthographic lexicon noting that short vowels are not spelled. Models of spoken word processing need to be revised with respect to the role of orthography in spoken word processing. Current models either ignore the role of orthography in spoken word processing altogether (McClelland & Elman, 1986; Marslen-Wilson & Tyler, 1980) or assume an obligatory (Grainger, Muneaux, Farioli, & Ziegler, 2005) or strategic (Norris, 1999) influence of orthography. The findings of this dissertation show how simultaneous presentation of phonological and orthographic (dis)similarities can influence word processing. They suggest that another theoretical account is also plausible and needs more attention in this field. Spoken words may not be processed based on two different independent sources of information i.e., phonology and orthography, but orthographic features may, in the course of development, restructure the lexicon such that both the phonological and the orthographic similarity between lexical entries determine their lexical distance. The orthographic effects leading to such a 'phonographic' organization of the mental lexicon seem to be largely involuntary.

Models of word processing assume interactive (e.g., (Grainger, Muneaux, Farioli, & Ziegler, 2005; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) or autonomous (Norris, 1999) processing levels. The outcome of this dissertation suggests that a dynamic model can best accommodate word processing. Word processing may undergo a developmental shift from interactive to autonomous processing with early interactions between phonology and orthography leading to the development of phonographic networks in the mental lexicon.

9-2. Suggestions for future research

The findings of this dissertation are limited to the research paradigms used to test the research hypotheses. It might be that other paradigms would yield different findings. The case of orthographically unrepresented short vowels in Persian lends itself well to investigate the influence of orthographic knowledge on speech processing at different levels. Many paradigms used in this dissertation tapped into the perceptual level. It is recommended to run more experiments that tap into lexical processing (e.g., auditory (primed) lexical decision tasks), to investigate whether orthographic knowledge for short or long vowels in Persian can influence processing words with short/long vowels at the lexical level.

References

Abu-Rabbia, S. (1997). Reading in Arabic Orthography: The Effect of Vowels and Context on Reading Accuracy of Poor and Skilled Native Arabic Readers in reading. *Journal of Psycholinuistic Research, 26(4)*, 465-482. doi:10.1023/a:1025034220924

Abu-Rabbia, S. (1998). Reading Arabic Texts: Effects of Text Type, Reader Type, and Vowelization. *Reading and Writing: An Interdesciplinary Journal, 10*, 106-119. doi: 10.1023/A:1007906222227

Adrian, J., Alegria, J., & Morais, J. (1995). Metaphonological Abilities of Spanish Illiterate Adults. *International Journal of Psychology, 30(3)*, 329-351. doi:10.1080/00207599508246574

Alario, F., Perre, L., Castle, C., & Ziegler, J. (2007). The Role of Orthograpgy in Speech Perception Revisited. *Cognition, 102(3)*, 464-475. doi: 10.1016/j.cognition.2006.02.002

Alegria, J. (1985). Por un enfoque psiolinguistico del aperendizaje de la lectura y sus dificultades (Support for a Psycholinguistic Approach to Reading Acquisition and Reading Difficulties: Twenty Years Later). *Journal for the Study of Education and Development, Infancia y Aprendizaje, 8*, 79-94. doi: 10.1080/02103702.1985.10822061

Alegria, J., Pignot, J., & Morais, J. (1982). Phonetic Analysis of Speech and Memory Codes in Beginning Readers. *Memory and Cognition, 10*, 451-456. doi:10.3758/BF03197647

Baddeley, A. (1979). Working Memory and Reading. In P. Kolers, M. Wrolstad, & H. Bouma (Eds.), *Processing of Visible Language* (pp. 355-370). New York: Plenum Press.

Baddeley, A. (1992). Working Memory. *Science, 255(5044)*, 556-559. doi: 10.1126/science.1736359

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annu. Rev. Psychol., 63*, 1-29. doi: 10.1146/annurev-psych-120710-100422

Bakhtiar, M., & Weekes, B. (2015). Lexico-Semantic Effects on Word Naming in Persian: Does Age of Acquisiton Have an Effect? *Memory and Cognition, 43(2)*, 298-313. doi: 10.3758/s13421-014-0472-4

Bakhtiar, M., Mokhlesin, M., Pattamadilok, C., Politzer-Ahles, S., & Zhang, C. (2021). The Effect of Orthographic Transparency on Auditory Word Recognition Across the Development of Reading Proficiency. *Frontiers in Psychology, 12*, 1-9. doi: 10.3389/fpsyg.2021.691989

Baluch, B. (1993). Lexical Decision in Persian: A Test of Orthographic Depth Hypothesis. *International Journal of Psychology, 28*, 19-29. doi: 10.1080/00207599308246915

Baluch, B. (2005). Persian Orthography and Its Relation to Literacy. In J. Malatesha & P. Aaron (Eds.), *Handbook of Orthography and Literacy* (pp. 365-376). London: Lawrence Erlbaum Associates.

Baluch, B., & Besner, D. (1991). Visual Word Recognition: Evidence for Strategic Control of Lexical and Nonlexical Routines in Oral Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 644-652. doi: 10.1037/0278-7393.17.4.644

Baluch, B., & Dana-ye-Tousie, M. (2006). Memory for Words as a Function of Spelling Transparency. *Journal of Psychology, 140*, 95– 104. doi: 10.3200/JRLP.140.2.95-104

Baluch, B., & Shahidi, S. (1991). Visual Word Recognition in Beginning Readers of Persian. *Perceptual and Motor Skills, 72(3)*, 1327-1331. doi: 10.2466/PMS.72.4.1327-1331

Barmes, D., Tager, I., Satariano, W., & Yaffe, K. (2004). The Relationship Between Literacy and Cognition in Well Educated Elders. *Journal of Gerontology: Medical Sciences, 59A*, 390-395. doi:10.1093/gerona/59.4.m390

Bassetti, B., Escudero, P., & Hayes-Harb, R. (2015). Second Language Phonology at the Interface between acoustic and Orthgraphc Input. *Applied Psycholinguistic, 36(1)*, 1-6. doi:10.1017/S0142716414000393

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*, 1-48. doi: 10.18637/jss.v067.i01

Battiston, F., Nicosia, V., & Latora, V. (2014). Structural Measures for Multiplex Networks. *Physical Review, 89(3)*. doi: 10.1103/PhysRevE.89.032804

Berent, I., & Perfetti, C. (1995). A Rose Is a Reez: The Two-Cycle Model of Phonology Assembly in Reading English. *102*, 146-184. doi: 10.1037/0033-295X.102.1.146

Bertelson, P., de Gelder, B., Tfouni, L., & Morais, J. (1989). Metaphonological Abilities of Adult Illiterates : New Evidence on Heterogeneity. *European Journal of Cognitive Psychology, 1(3)*, 239-250. doi: 10.1080/09541448908403083

Bialystok, E. (1986). Factors in the Growth of Linguistic Awareness. *Child Dvelopment, 57*, 498-510. doi: 10.2307/1130604

Boersma, P., & Weenink, D. (2015). *Praat: Doing Phonetics by Computer*. https://www.fon.hum.uva.nl/praat/

Boudelaa, S., & Marslen-Wilson, W. (2015). Structure, Form, and Meaning in the Mental Lexicon: Evidence from Arabic. *Language, Cognition, and Neusoscience, 30(8)*, 955-992. doi: 10.1080/23273798.2015.1048258

Broersma, M. (2005). Perception of Familiar Contrasts in Unfamiliar Position. *Journal of the Acoustical Society of America, 117(6)*, 3890-3901. doi:10.1121/1.1906060

Bruce, D. (1964). The Analysis of Word Sounds By Young Children. *British Journal of Educational Psychology, 34(2)*, 158-170. doi:https://doi.org/10.1111/j.2044-8279.1964.tb00620.x

Bryant, P., & Bradley, L. (1985). *Children's Reading Problems: Psychology and Education.* Oxford: University of Illinois Press.

Bryant, P., Maclean, M., Bradley, L., & Crossland, M. (1990). Rhyme, Alliteration, Phoneme Detection, and Learning to Read. *Developmental Psychology, 26(3)*, 158-170. doi:10.1037/0012-1649.26.3.429

Calfee, R. (1977). Assessment of Individual Reading Skills: Basic Research and Practical Applications. In A. Reber & D. Scarborough (Eds.), *Toward a Psychology of Reading* (pp. 289-323). Hillsdale, NJ: Earlbaum.

Campbell, G., & King, G. (1991). *Compendium of the World's Languages.* New York: Routledge.

Caravolas, M., & Bruck, M. (1993). The Effect of Oral and Written Language Input in Children's Phonological Awareness: A Cross Linguistic Study. *Journal of experimental child Psychology, 50*, 1-30. doi: 10.1006/jecp.1993.1001

Carrillo, M. (1994). Development of Phonological Awareness and Reading Acquisition: A Study in Spanish Language. *Reading and Writing, 6*, 279-298. doi:10.1007/BF01027086

Cary, L. (1988). A Anälise Explfcita das Unidades da Fala nos Adultos Näo-alphabetizados. *PhD Dissertation*. Portugal: University of Lisbon. http://hdl.handle.net/10451/42414

Cassar, M., & Treiman, R. (1997). The Beginnings of Orthographic Knowledge: Children's Knowledge of Double Letters in Words. *Journal of Educational Psychology, 89(4)*, 631-644.

Castle, A. (2006). The Dual Route Model and the Developmental Dyslexias. *London Review of Education, 4*, 49-61. doi: 10.1080/13603110600574454

Castle, A., & Coltheart, M. (2004). Is There a Causal Link from Phonological Awareness to Success in Learning to Read? *Cognition, 91(1)*, 77-111. doi:10.1016/s0010-0277(03)00164-1

Castle, A., Holmes, V., Neath, J., & Kinoshita, S. (2003). How Does Orthographic Knowledge Influence Performance on Phonological Awareness Tasks? *Quarterly Journal of Experimental Psychology, 56*, 445-467. doi:0.1080/02724980244000486

Chen, J., Chen, T., & Dell, J. (2002). Word-Form Encoding in Mandarin Chinese as Assessed by the Implicit Priming Task. *Journal of Memory and Language, 46(4)*, 751-781. doi: 10.1006/jmla.2001.2825

Chereau, C., Gaskell, M. G., & Dumay, N. (2007). Reading Spoken Words: Orthographic Effects in Auditory Priming. *Cognition, 102(3)*, 341-360. doi: 10.1016/j.cognition.2006.01.001

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English.* N.Y.: Harper and Row.

Clifton, C. (2015). The Role of Phonology in Reading. A Selective Review. In L. Frazier & E. Gibson (Eds.), *Explicit and Implicit Prosody in Sentence Processing* (pp. 161-176). New York: Springer.

Coltheart, M. (1978). Lexical Access in Simple Reading Tasks. In G. Underwood (Ed.), *Strategies of Information Processing* (pp. 151-216). San Diego, CA: Academic Press.

Coltheart, M. (2005). Modeling Reading: The Dual-Route Approach. In M. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 6-23). Blackwell Publishing Ltd.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review, 108*, 204-256. doi:10.1037/0033-295x.108.1.204

Content, A. (1984). L'analyse phonétique explicite de la parole et l'acquisition de la lecture. *L'Année Psychologique, 84*, 555-572. doi:10.3406/psy.1984.29053

Cortese, M., & Simpson, G. (2000). Regularity Effects in Word Naming. What are They? *Memory & Cognition, 28(8)*, 1269-1276. doi: 10.3758/BF03211827

Cossu, G., Shankweiler, D., Liberman, I., Katz, l., & Tola, G. (1988). Awareness of Phonological Segments and Reading Ability in Italian Children. *Applied psycholinguistcs, 9*, 1-16. doi:10.1017/S0142716400000424

Crowder, R. (1982). Decay of Auditory Memory in Vowel Discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 153-162. doi:10.1037//0278-7393.8.2.153

Cutler, A., & Davis, C. (2012). An Orthographic Effect in Phoneme Processing, and Its Limitations. *Frontiers in Psychology, 3*, 1-7. doi: 10.3389/fpsyg.2012.00018

Cutler, A., & Norris, D. (1979). Monitoring sentence comprehension. In W. Cooper & E. Walker (Eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett* (pp. 113–134). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme Identification and the Lexicon. *Cognitive Psychology, 19*, 141-177. doi: 10.1016/0010-0285(87)90010-7

Cutler, A., Norris, D., & Van Ooijen, B. (1990). Vowels as Phoneme Detection Targets. *Proceedings of the First International Conference on spoken Language Processing*

(ICSLP 1990), 581-584. https://www.mpi.nl/publications/item77007/vowels-phoneme-detection-targets

Cutler, A., Treiman, R., & Van Ooijen, B. (2010). Strategic Deployment of Orthographic Knowledge in Phoneme Detection. *Language and Speech, 53*, 307-320. doi:10.1177/0023830910371445

Cutler, A., Van Ooijen, B., Norris, D., & Sanchez-Casas, R. (1996). Speeded Detection of Vowels: A Cross Linguistic Study. *Perception and Psychophysics, 58*, 807-822. doi:10.3758/bf03205485

Damian, M., & Bowers, J. (2003). Effects of Orthography on Speech Production in a Form Preparation Paradigm. *Journal of Memory and Language, 49(1)*, 119-132. doi: 10.1016/S0749-596X(03)00008-1

Damian, M., & Bowers, J. (2009). Assessing the Role of Orthography in Speech Perception and Production: Evidence from Picture-Word Interference. *Euoropean Journal of Cognitive Psychology, 21(4)*, 581-598. doi: 10.1080/09541440801896007

Daniels, P., & Bright, W. (1996). *The World's Writing Systems.* New York: Oxford University Press.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The Neural Code for Written Words: A Proposal. *Trends in Cognitive Science, 9*, 335-341. doi: 10.1016/j.tics.2005.05.004

Deo, A., & Kiparsky, P. (2011). Poetries in Contact: Arabic, Persian, and Urdu. In M.K. Lotman and M. Lotman (Eds.), *Proceedings of International Conference on Frontiers in Comparative Metrics* (pp. 147–173). Estonia.

Diependaele, K., Ziegler, J., & Grainger, J. (2010). Fast Phonology and the Bi-Modal Interactive-Activation Model. E*uropean Journal of Cognitive Psychology, 22*, 764–778. doi:10.1080/09541440902834782

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic Effects on Phoneme Monitoring. *Canadian Journal of Experimental Psychology, 49(2)*, 264-271. doi:10.1037/1196-1961.49.2.264

Donnenwerth-Nolan, S., Tanenhaus, M., & Seidenberg, M. (1981). Multiple Code Activation in Word Recognition: Evidence from Rhyme Monitoring. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 170-180. doi:10.1037/0278-7393.7.3.170

Dufour, S., & Peerman, R. (2004). Phonological Priming in Auditory Word Recognition: Initial Overlap facilitation Effect Varies as a Function of Target Word Frequecy. *Behavior, Brain, and Cognition, 3*, 1-12. doi: 10.4000/cpl.437

Dupoux, E. (1993). The Time course of Prelexical Processing: The Syllabic Hypothesis Revisited. In G. Altmann & R. Shillcock (Eds.), *Cognitive Models of Speech Processing: The Second Sperlonga Meeting* (pp. 81–114). UK: Psychology Press.

Ehri, L. (1984). How Orthography Alters Spoken Language Competencies in Children Learning to Read and Spell. In J. Downing & R. Valtin (Eds.), *Language awareness and Learning to Read* (pp. 119-147). New York: Springer Verlag. doi: 10.1007/978-1-4613-8248-5_7

Ehri, L. (1985). Effects of Printed Language Acquisition on Speech. In D. Olson, N. Torrance, & A. Hildyard (Eds.), *Literacy, Language, and Learning* (pp. 333-367). London, New York, New Rochelle, Melborne, Sydney: Cambridge University Press.

Ehri, L. (1987). Learning to Read and Spell Words. *Journal of Reading Behavior, 19(1)*, 5-31. doi: 10.1080/10862968709547585

Ehri, L. (2005). Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading, 9(2)*, 167-188. doi: 10.1207/s1532799xssr0902_4

Ehri, L. (2014). Orthographic Mapping in the Acquisition of Sight Word Reading, Spelling Memory, and Vocabulary Learning. *Scientific Studies of Reading, 18(1)*, 5-21. doi: 10.1080/10888438.2013.819356

Ehri, L., & Wilce, L. (1979). The Mnemonic Value of Orthography among Beginning Readers. *Journal of Educational Psychology, 71(1)*, 26-40. doi: 10.1037/0022-0663.71.1.26

Ehri, L., & Wilce, L. (1980). The Influence of Orthography on Readers' Conceptualization of the Phonemic Structure of Words. *Applied Psycholinguistics, 1(4)*, 371-385. doi: 10.1017/S0142716400009802

Escudero, P., & Wanrooij, K. (2010). The Effect of L1 Orthography on Non-native Vowel Perception. *Lang. Speech, 53(3)*, 343-365. doi: 10.1177/0023830910371447

Escudero, p., Boersma, P., & Simon, E. (2013). Learning Words in a Third Language: Effects of Vowel Inventory and Language Proficiency. *Language and Cognitive Processes, 28(6)*, 746-761. doi: 10.1080/01690965.2012.662279

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel Second-Language Words and Asymmetric Lexical Access. J*ournal of Phonnetics, 36(2)*, 345-360. doi: 10.1016/j.wocn.2007.11.002

Esfahbod, B. (2012, February 2). *Persian Word Frequency*. Retrieved at 2019, Aug. 16 from https://github.com/behnam/persian-word-frequency

Federmeier, K., Kutas, M., & Dickson, D. (2016). A Common Neural Progression to Meaning in about a Third of a Second. In G. Hickok & S. Small (Eds.), *Neurobiology of Language* (pp. 557-567). Elsevier Inc. doi:10.1016/B978-0-12-407794-2.00045-6

Ferrand, L. (2000). Reading Aloud Polysyllabic Words and Nonwords: The Syllabic Length Effect Reexamined. *Psychonomic Bulletin & Review, 7*, 142-148. doi: 10.3758/BF03210733

Ferrand, L., & Grainger, J. (2003). Homophone Interference Effects in Visual Word Recognition. *The Quarterly Journal of Experimental Psychology, 56*, 403-419. doi:10.1080/02724980244000422

Fodor, J. (1983). *The Modularity of Mind.* Cambridge, MA: MIT Press.

Foss, D., & Blank, M. (1980). Identifying the Speech Codes. *Cognitive Psychology, 12(1)*, 1–31. doi: 10.1016/0010-0285(80)90002-X

Foss, D., & Gernsbacher, M. (1983). Cracking the Dual Code: Toward a Unitary Model of Phoneme Identification. *Journal of Verbal Learning and Verbal Behavior, 22(6)*, 609–632. doi: 10.1016/S0022-5371(83)90365-1

Foss, D., & Swinney, D. (1973). On the Psychological Reality of the Phoneme: Perception, Identification, and Consciousness. *Journal of Verbal Learning and Verbal Behavior, 12(3)*, 246–257. doi: 10.1016/S0022-5371(73)80069-6

Fowler, A. (1991). How Early Phonological Development Might Set the Stage for Phoneme Awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological Processes in Literacy* (pp. 97–117). Hillsdale, NJ: Erlbaum.

Frauenfelder, U., & Segui, J. (1989). Phoneme Monitoring and Lexical Processing: Evidence for Associative Context Effects. *Memory & Cognition, 17*, 134–140. doi: 10.3758/BF03197063

Frauenfelder, U., Segui, J., & Dijkstra, T. (1990). Lexical Effects in Phonemic Processing: Facilitatory or Inhibitory? *Journal of Experimental Psychology: Human Perception and Performance, 16(1)*, 77–91. doi: 10.1037/0096-1523.16.1.77

Frith, U. (1998). Literally Changing the Brain. *Brain, 121*, 1010-1012. doi: 10.1093/brain/121.6.1011

Frost, R. (1995). Phonological Computation and Missing Vowels: Mapping Lexical Involvement in Reading. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21(2)*, 398-408. doi: 10.1037//0278-7393.21.2.398

Frost, R. (1998). Toward a Strong Phonological Theory of Visual Word Recognition: True Issues and False Trails. *Psychol Bull, 123(1)*, 71-99. doi: 10.1037/0033-2909.123.1.71

Frost, R. (2005). Orthographic Systems and Skilled Word Recognition Processes in Reading. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 272-295). Malden, MA: Blackwell Publishing. doi:10.1002/9780470757642.ch15

Gelb, I. (1952). *A Study of Writing.* Chicago: University of Chicago Press.

Gentner, D. (1983). Structure-mapping: A Theoretical Framework for Analogy. *Cognitive Scienece, 7(2),* 155-170. doi: 10.1016/S0364-0213(83)80009-3

Gerrits, E., & Schouten, M. (2004). Categorical Perception Depends on the Discrimination Task. *Perception and Psycholinguistics, 66*, 363-376. doi: /10.3758/BF03194885

Goldblum, M., & Matute de Duran, E. (2000). Are Illiterate People Deep Dyslexic? *Journal of Neurolinguistics, 2*, 103-114. doi: 10.1016/S0911-6044(86)80006-9

Gombert, J. (1992). *Metalinguistic Development.* Chicago: University of Chicago Press.

Gonzalez, J., & Gonzalez, M. (1993). Phonological Awareness in Learning Literacy. *Cognition, 5(2)*, 153-170. doi: 10.3406/intel.1994.1454

González-Valenzuela, M., Díaz-Giráldez, F., & López-Montiel, M. (2016). Cognitive Predictors of Word and Pseudoword Reading in Spanish First-Grade Children Spanish First-Grade Children. *Front. Psychol.*, 7, 774. doi: 10.3389/fpsyg.2016.00774

Gordon, M. (2004). Syllable Weight. In B. Hayes, R. Kirchner, & D. Steriade, *Phonetic Bases for Phonological Markedness.* Cambridge: Cambridge University Press.

Goswami, U. (2000). Phonological Representations: Reading Development and Dyslexia: Towards a Cross-Linguistic Theoretical Framework. *Dyslexia*, *6(2)*, 133-151. doi: 10.1002/(SICI)1099-0909(200004/06)6:2<133::AID-DYS160>3.0.CO;2-A

Goswami, U., & Bryant, P. (1990). *Phonological Skills and Learning to Read.* Hillsdale, NJ: Erlbaum.

Grainger, J., & Ferrand, L. (1996). Masked Orthographic and Phonological Priming in Visual Word Recognition and Naming: Cross-task Comparisons. *Journal of Memory and Language, 53(5)*, 623-647. doi: 10.1006/jmla.1996.0033

Grainger, J., Muneaux, M., Farioli, F., & Ziegler, J. (2005). Effects of Phonological and Orthographic Neighbourhood Density Interact in Visual Word Recognition. *Quarterly Journal of experimental Psychology*, *58(6)*, 981–998.doi: 10.1080/02724980443000386

Hahn, U., & Bailey, T. (2005). What Makes Words Sound Simlar? *Cognition*, *97(3)*, 227-267. doi: 10.1016/j.cognition.2004.09.006

Halle, M., & Vergnaud, J. (1980). Three Dimensional Phonology. *Journal of Linguistic Research I*, 83-105. http://web.mit.edu/morrishalle/pubworks/papers/1980_Halle_Vergnaud_Three-dimensional_phonology.pdf

Halle, P., Chereau, C., & Segui, J. (2000). Where Is the /b/ in "Absurde" [apsyrd]? It Is in French Listeners' Minds. *Journal of Memory & Language, 43(4)*, 618-639. doi: 10.1006/jmla.2000.2718

Hayden, C. (2018, November 16). *The Library of Congress*. Retrieved at 2019, February 25 from https://www.loc.gov/exhibits/thousand-years-of-the-persian-book/writing-systems-and-scripts.html

Hayes, B. (1989). Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry, 20(2)*, 253-306. http://www.jstor.org/stable/4178626

Hebb, D. (1949). *The Organization of Behavior: A Neuropsychological Theory.* NewYork: John Wiley & Sons.

Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2011). Categorical Perception of Voicing, Colors, and Facial Expressions: A Developmental Study. *Speech Communication*, *53(3)*, 417-430. doi: 10.1016/j.specom.2010.11.005

Hosseini Almadani, F. (2016). The Effect of Script on Reading Development of Farsi and Arabic Bilinguals. *Master Thesis*. Wilfrid Laurier University.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal, 50*, 346-363. doi: 10.1002/bimj.200810425

Hulme, C., Caravolas, M., Malkova, G., & Brigstocke, S. (2005). Phoneme Isolation Ability Is Not Simply a Consequence of Letter-Sound Knowledge. *Cognition, 97(1)*, B1-B11. doi: 10.1016/j.cognition.2005.01.002

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0 [Software]. Armonk, NY: IBM Corp.

Indefrey, P., & Levelt, W. (2004). The Spatial and Temporal Signatures of Word Production Components. *Cognition, 92,* 101-144. doi: 10.1016/j.cognition.2002.06.001

Jackson, N., & Coltheart, M. (2001). *Routes to Reading Success and Failure: Toward an Integrated Cognitive Psychology of Atypical Reading.* Psychology Press.

Jakimik, J., Cole, R., & Rudnicky, A. (1985). Sound and Spelling in Spoken Word Recognition. *Journal of Memory and Language, 24*, 165-178. doi: 10.1016/0749-596X(85)90022-1

Juphard, A., Carbonnel, S., Ans, B., & Valdois, S. (2006). Length Effect in Naming and Lexical Decision: The Multitrace Memory Models's Account. *Current Psychology Letters: Behavior, Brain and Cognition, 19(2)*, 1-14. doi: 10.4000/cpl.1005

Kahnemuyipour, A. (2003). Syntactic Categories and Persian Stress. *Natural Language & Linguistic Theory, 21*, 333-379. doi: 10.1023/A:1023330609827

Karmiloff-Smith, A. (1995). *Beyond Modularity: A Developmental Perspective on Cognitive Science.* The MIT Press.

Katz, L., & Frost, R. (1992). The Reading Process Is Different for Different Orthographies: The Orthographic Depth Hypothesis. In R. Frost & L. Katz (Eds.), *Advances in Psychology* (pp. 67-84). North-Holland. doi:10.1016/S0166-4115(08)62789-2

Katz, L., Brancazio, L., Irwin, J., Katz, S., Magnuson, J., & Whalen, D. (2012). What Lexical Decision and Naming Tell Us about Reading. *Reading and Writing, 25*, 1259–1282. doi: 10.1007/s11145-011-9316-9

Kello, C., & Beltz, B. (2009). Scale-free Networks in Phonological and Orthographic Wordform Lexicons. In I. Chitoran, C. Coupé, E. Marsico, & F. Pellegrino, *Approaches to Phonological Complexity* (pp. 171-190). Berlin: Mouton de Gruyter. doi: 10.1515/9783110223958.171

Kolinsky, R. (1998). Spoken Word Recognition: A Stage-processing Approach of Language Differences. *European Journal of Cognitive Psychology, 10*, 1-40. doi: 10.1080/713752266

Kolinsky, R. (2015). How Learning to Read Influences Language and Cognition. In A. Pollatsek & R. Treiman (Eds.), *The Oxford Handbook of Reading* (pp. 377–393). Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199324576.013.29

Kolinsky, R., Morais, J., Content, A., & Carry, L. (1987). Finding Parts within Figures: A Developmental Study. *Perception, 16(3)*, 399–407. doi: 10.1068/p160399

Kolinsky, R., Navas, A., Vidigal de Paula, F., Riberio de Brito, N., de Medeiros Botecchia, L., Bouton, S., & Serniclaes, W. (2021). The Impact of Alphabetic Literacy on the Perception of Speech Sounds. *Cognition, 213,* 1-12. doi:10.1016/j.cognition.2021.104687

Koriat, A. (1984). Reading Without Vowels: Lexical Access in Hebrew. In H. Bouma & D. Bouwuis (Eds.), *Attention and Performance: Control of Language Processes* (pp. 227-274). London: Lawerence Erlbaum Associates.

Koriat, A. (1985). lexical Access for Low and High Frequency Words in Hebrew. *Mamory and Cognition, 13*, 37-44. doi: 10.3758/BF03198441

Kramsky, J. (1966). Some Remarks on the Problem of Quantity of Vowel Phonemes in Modern Persian. *Archiv Orientalni, 34*, 215-220. https://www.proquest.com/scholarly-journals/some-remarks-problem-quantity-vowel-phonemes/docview/1304096586/se-2?accountid=14532

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82(13),* 1-26. doi: 10.18637/jss.v082.i13

Kwok, R., Cuetos, F., Avdyli, R., & Ellis, W. (2006). Reading and Lexicalization in Opaque and Transparent Orthographies: Word Naming and Word Learning in English and Spanish. *Quarterly Journal of Experimental Psychology, 70(10)*, 1-57. doi: 10.1080/17470218.2016.1223705

Lazard, G. (1992). *Grammar of Contemporary Persian.* California: Shirley Lyons.

Leinenger, M. (2014). Phonological Decoding during Reading. *Psycol. Bull., 140*, 1534-1555. doi: 10.1037/a0037830

Levelt, W. (1989). *Speaking: From Intention to Articulation.* Cambridge, MA: MIT Press.

Lewis, M. (2009). *Ethnologue: Languages of the World.* Dallas, Tex.: SIL International.

Liberman, I., & Shankweiler, D. (1977). Speech, the Alphabet and Teaching to Read. In L. Resnicky & P. Weaver (Eds.), *Theory and Practice of Early Reading* (pp. 105-129). Hillsdale, NJ: Erlbaum.

Liberman, I., Shankweiler, D., Fischer, F., & Carter, B. (1974). Explicit Syllable and Phoneme Segmentation in the Young Child. *Journal of Experimental Child Psychology, 18(2)*, 201–212. doi: 10.1016/0022-0965(74)90101-5

Liberman, I. (1973). Segmentation of the Spoken Word and Reading Acquisition. *Bulletin of Orton Society, 23*, 65-77. doi: 10.1007/BF02653842

Linell, P. (1979). *Psycholgical Reality in Phonology.* Cambridge: Cambridge University Press.

Lukatela, G., & Turvey, M. T. (1994). Visual Lexical Access Is Initially Phonological: Evidence from Associative Priming by Words, Homophones, and Pseudohomophones. *Journal of Experimental Psychology, 123(2),* 107-128. doi:10.1037/0096-3445.123.2.107

Lukatela, G., Frost, S. J., & Turvey, M. T. (1998). Phonological Priming by Masked Nonword Primes in the Lexical Decision Task. *Journal of Memory and Language, 39*, 666–683. doi:10.1006/jmla.1998.2599

Lukatela, K., Carello, C., Shankweiler, D., & Liberman, I. (1995). Phonological Awareness in Illiterates: Observations from Serbo-Croatian. *Applied Psycholinguistics, 16(4)*, 463-487. doi: 10.1017/S0142716400007487

Lundberg, I., & Hoien, T. (1991). Initial Enabling Knowledge and Skills in Reading Acquisition: Print Awareness and Phonological Segmentation. In D. Sawyer & B. Fox (Eds.), *Phonological Awareness in Reading. The Evolution of Current Perspective* (pp. 73-95). New York: Springer-Verlag. doi: 10.1007/978-1-4612-3010-6_3

Lupker, S., & Williams, B. (1988). Rhyme Priming of Pictures and Words: A Lexical Activation Account. *Journal of Experimental Psychology*, 15(6), 1033-1046. doi: 10.1037/0278-7393.15.6.1033

Maclean, M., Bryant, P., & Bradley, L. (1987). Rhyme, Nursery Rhymes, and Reading in Early Childhood. *Merrill-Palmer Quarterly*, *33(3)*, 255-282. http://www.jstor.org/stable/23086536

Mann, V. (1986). Phonological Awareness: The Role of Reading Experience. *Cognition*, *24*, 65-92. doi: 10.1016/0010-0277(86)90005-3

Mann, V. (1991). Phonological Awareness and Early Reading Ability: One Perspective. In D. Sawyer & B. Fox (Eds.), *Phonological Awareness in Reading. The Evolution of Current Perspective* (pp. 191-215). New York: Springer-Verlag.

Mann, V., & Ditunno, P. (1990). Phonological Deficiencies: Effective Predictors of Future Reading Problems. In G. Pavlides (Ed.), *Dyslexia: A Neuropsychological and Learning Perspective* (pp. 105-131). Sussex, England: J, Wiley and Sons.

Mann, V., & Wimmer, H. (2002). Phoneme Awareness and Pathways into Literacy: A Comparison of German and American Children. *Reading and Writing, 15*, 653-682. doi: 10.1023/A:1020984704781

Markman, A., & Gentner, D. (1993). Structural Alignment during Similarity Comparisons. *Cognitive Psychology*, *25*, 431-467. doi: 10.3758/BF03212423

Markman, A., & Gentner, D. (1996). Commonalities and Differences in Similarity Comparisons. *Memory and Cognition*, *24*, 235-249. doi: 10.3758/BF03200884

Marslen-Wilson, W. (1984). Function and Process in Spoken Word Recognition: A Tutorial Review. In H. Bouma & D. Bouwhuis (Eds.), *Attention and Performance* (pp. 125-150). Lawrence Erlbaum Associates.

Marslen-Wilson, W., & Tyler, L. (1980). The Temporal Structure of Spoken Language Understanding. *Cognition, 8*, 1-71. doi: 10.1016/0010-0277(80)90015-3

Marslen-Wilson, W., & Welsh, A. (1978). Processing Interactions and Lexical Access during Word Recognition in Contineous Speech. *Cognitive Psychology, 10*, 29-63. doi: 10.1016/0010-0285(78)90018-X

Martin, R., & Jensen, C. (1988). Phonological Priming in the Lexical Decision Task: A Failure to Replicate. *Memory and Cognition, 16*, 505-521. doi: 10.3758/BF03197052

Massaro, D., & Jesse, A. (2005). The Magic of Reading: Too Many Influences for Quick and Easy Explanations. In T. Trabasso, J. Sabatini, D. Massaro, & R. Calfee (Eds.), *From orthography to pedagogy: Essays in honor of Richard L. Venezky* (pp. 37–61). Lawrence Erlbaum Associates Publishers.

McClelland, J., & Elman, J. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology, 18(1)*, 1-86. doi: 10.1016/0010-0285(86)90015-0

McClelland, J., & Rumelhart, D. (1981). An Interactive Activation Model of Context Effects in Letter: Part 1. An Account of Basic Findings. *Psychological Review, 88*, 375-407. doi: 10.1037/0033-295X.88.5.375

McClelland, J., Mirman, D., & Holt, L. L. (2006). Are There Interactive Processes in Speech Perception? *Trends Cogn Sci, 10(8)*, 363–369. doi:10.1016/j.tics.2006.06.007

McClelland, J. (1987). The Case of Interactionism in Language Processing. In M. Coltheart (Ed.), *Attention and Performance 12: The Psychology of Reading* (pp. 3–36). Lawrence Erlbaum Associates, Inc.

McQueen, J., Cutler, A., & Norris, D. (2000). Why MERGE Really Is Autonomous and Parsimonious. In A. Cutler, M. McQueen, & R. Zondervan (Eds.), *Proceedings of SWAP (Workshop on Spoken Word Access Processes)* (pp. 47-50). Nijmegen: Max-Planck-Institute for Psycholinguistics.

Mehler, J., Dommergues, J., Frauenfelder, U., & Segui, J. (1981). The Syllables Role in Speech Segmentation. *Journal of Verbal Leraning and Verbal Behavior, 20(3)*, 298-305. doi: 10.1016/S0022-5371(81)90450-3

Meyer, A. (1990). The Time Course of Phonological Encoding in Language Production: The Encoding of Successive Syllable. *Journal of Memory and Language, 29(5)*, 524-545. doi: 10.1016/0749-596X(90)90050-A

Meyer, D., & Schwaneeldt, R. (1971). Facilitation in Recognizing Pairs of words: Evidence of a Depedence between Retrieval Operations. *Journal of Experimental Psychology*, *90*, 227-234. doi: 10.1037/h0031564

Meyer, D., Schvaneveldt, R., & Ruddy, M. (1974). Functions of Graphemic and Phonemic Codes in Visual Word Recognition. *Memory and Cognition, 2*, 309-321. doi: 10.3758/BF03209002

Mo'in, M., & Shahidi, J. (1972). *Mo'in Encyclopedic Dictionary.* Tehran: Amirkabir.

Monsell, S., Doyle, M., & Haggard, P. (1989). Effects of Frequency on Visual Word Recognition Tasks: Where Are They? *Journal of Experimental Psychology, 118*, 43-71. doi: 10.1037//0096-3445.118.1.43

Morais, J. (1991). Phonological Awareness: A Bridge between Language and Literacy. In D. Sawyer & B. Fox (Eds.), *Phonological Awareness in Reading. The Evolution of Current Perspective* (Vol. 28, pp. 31-71). New York: Springer-Verlag. doi: 10.1007/978-1-4612-3010-6_2

Morais, J. (2003). Levels of Phonological Representation in Skilled Reading and in Learning to Read. *Reading and Writing: An Interdisciplinary Journal, 16*, 123-151. doi: 10.1023/A:1021702307703

Morais, J. (2021). The Phoneme: A Conceptual Heritage from Alphabetic Literacy. C*ognition, 213.* doi: 10.1016/j.cognition.2021.104740

Morais, J., & Kolinsky, R. (1994). Perception and Awareness in Phonological Processing: The Case of the Phoneme. *Cognition, 50*, 287-297. doi: 10.1016/0010-0277(94)90032-9

Morais, J., & Kolinsky, R. (1995). The Consequences of Phonemic Awareness. In B. de Gelder & J. Morais (Eds.), *Speech and Reading: Comparative Approaches* (pp. 317–337). Hove, UK: Psychology Press.

Morais, J., & Kolinsky, R. (2002). Literacy Effects on Language and Cognition. In L. Bäckman & C. von Hofsten (Eds.), *Psychology at the Turn of the Millennium, Cognitive, Biological, and Health Perspectives* (pp. 507–530). UK: Psychology Press.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy Training and Speech Segmentation. *Cognition, 24*, 45-64. doi: 10.1016/0010-0277(86)90004-1

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does Awareness of Speech as a Sequence of Phones Arise Spontaneously? *Cognition, 7(4)*, 323-331. doi: 10.1016/0010-0277(79)90020-9

Morais, J., Castro, S., Scliar-Cabral, L., Kolinsky, R., & Content, A. (1987). The Effect of Literacy on the Recognition of Dichotic Words. *Quarterly Journal of Experimental Psychology, 39(3)*, 451-465. doi: 10.1080/14640748708401798

Morais, J., Content, A., Bertelson, P., Cary, L., & Kolinsky, R. (1988). Is There a Sensitive Period for the Acquisition of Segmental Analysis? *Cognitive Neuropsychology, 5*, 347-352. doi: 10.1080/02643298808252940

Morais, J., Kolinsky, R., & Claes, T. (1995). On the Relations between Phonological and Metaphonological Processing: A Study Using the ABX Discrimination Task. In C. Sorin & M. Wajskop (Eds.), *Levels in Speech Communication: Relations and Interactions* (pp. 201-210). Amsterdam: Elsevier Science B.V.

Mroczko-Wasowicz, A., & Nikolic, D. (2014). Semantic Mechanisms May be Responsible for Developing Synesthesia. *8, 509*. doi:10.3389/fnhum.2014.00509

Muneaux, M., & Ziegler, J. (2004). Locus of Orthographic Effects in Spoken Word Recognition: Novel Insights from the Neighbour Generation Task. *Language and Cognitive Processes, 19*, 641-660. doi: 10.1080/01690960444000052

Nakamura, M., Kolinsky, R., Spagnoletti, C., & Morais, J. (1998). Phonemic Awareness in Alphabetically Literate Japanese Adults: The Influence of the First Acquired Writing

System. *Cahiers de Psychologie Cogrffive/Current Psychology of Cognition, 17(2)*, 417-450. https://docslib.org/doc/8364811/phonemic-awareness-in-alphabetically-literate

Natel Khanlari, P. (1987). *Tarikh-e Zaban-e Farsi (Persian Language History).* Tehran: Nashr-e No.

Nation, K., & Hulme, C. (2011). Learning to Read Changes Children's Phonological Skills: Evidence from a Latent Variable Longitudinal Study of Reading and Nonword Repetition. *Developmental Science, 14*, 649-659. doi: 10.1111/j.1467-7687.2010.01008.x

Navon, D., & Shimron, J. (1981). Does Word Naming Involve Grapheme to Phoneme Translation? Evidence From Hebrew. *Journal of Verbal Learning and Verbal Behavior, 20(1)*, 97-109. doi: 10.1016/S0022-5371(81)90334-0

Navon, D., & Shimron, J. (1985). *Source and Function of Phonological Codes in Processing Written Hebrew: Final Report to the US-Israel Binational Science Foundation.* Haifa, Israel: University of Haifa Press.

Nayernia, L., van de Vijver, R., & Indefrey, P. (2019). The Influence of Orthography on Phonemic Knowledge: An Experimental Investigation on German and Persian. *Journal of Psycholinguistic Research, 48*, 1391-1406. doi:10.1007/s10936-019-09664-9

Nix, A., Mehta, G., Dye, J., & Cutler, A. (1993). Phoneme Detection as a Tool for Comparing Perception of Natural and Synthetic Speech. *Computer Speech and Language*, *7*, 211–228. doi: 10.1006/csla.1993.1011

Norris, D. (1999). The Merge Model: Speech Perception Is Bottom-up. *The Journal of the Acoustical Society of America*, *106(4)*, 2295. doi: 10.1121/1.427854

Norris, D., McQueen, J., & Cutler, A. (2000). Merging Information in Speech Recognition. Feedback Is Never Necessary. *Behavioral and Brain Science, 23*, 299-325. doi:10.1017/S0140525X00003241

Norris, D., McQueen, J., & Cutler, A. (2003). Perceptual Learning in Speech. *Cognitive Psychology, 47*, 204-238. doi: 10.1016/s0010-0285(03)00006-9

Olson, D. (1998). *The World on Paper: The Conceptual and Cognitive Implications of Writing and Reading.* Cambridge: Cambridge University Press.

Ostrosky-Solis, F., Ardila, A., & Rosselli, M. (1999). Neuropsi: A Brief Neuropsychological Test Battery in Spanish with Norms by Age and Educational Level. *Journal of the International Neuropsychological Society, 5(5)*, 413-433. doi: 10.1017/S1355617799555045

Pattamadilok, C., Knierim, I., Kawabata Duncan, K., & Devlin, J. (2010). How Does Learning to Read Affect Speech Perception? *The Journal of Neurosceince, 30(25)*, 8435-8444. doi: 10.1523/JNEUROSCI.5791-09.2010

Pattamadilok, C., Kolinsky, R., Luksaneeyanawin, S., & Morais, J. (2008). Orthographic Congruency Effects in the Suprasegmental Domain: Evidence from Thai. *Quarterly Journal of Experimental Psychology*, *61(10),* 1515-1537. doi: 10.1080/17470210701587305

Pattamadilok, C., Kolinsky, R., Ventura, P., Radeau, M., & Morais, J. (2007). Orthographic Representations in Spoken Word Priming: No Early Automatic Activation. *Language and Speech, 50,* 505-531. doi: 10.1177/00238309070500040201

Pattamadilok, C., Lafontaine, H., Morais, J., & Kolinsky, R. (2010). Auditory Word Serial Recall Benefits from Orthographic Dissimilarity. *Language and Speech, 53(3)*, 321-341. doi: 10.1177/0023830910371450

Pattamadilok, C., Morais, J., & Kolinsky, R. (2011). Naming in Noise: The Contribution of Orthographic Knoweldge to Speech Repetition. *Front. Psychol., 2*, 361. doi: 10.3389/fpsyg.2011.00361

Pattamadilok, C., Morais, J., De Vylder, O., Ventura, P., & Kolinsky, R. (2009). The Orthographic Consistency Effect in the Recognition of French Spoken Words: An Early Development Shift from Sublexical to Lexical Orthographic Activation. *Appl. Psycholinguist., 30*, 441-462. doi:10.1017/s0142716409090225

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The Locus of the Orthographic Consistency Effect in Auditory Word Recognition: Further Evidence from French. *Language and Cognitive Processes, 22(5)*, 700-726. doi: 10.1080/01690960601049628

Pattamadilok, C., Nelis, A., & Kolinsky, R. (2013). How Does Reading Performance Modulate the Impact of Orthographic Knowledge on Speech Processing? A Comparison of Normal Readers and Dyslexic Adults. *Annals of Dyslexia, 64(1)*, 57-76. doi: 10.1007/s11881-013-0086-8

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. (2009). On-line Orthographic Influences on Spoken Language in a Semantic Task. *Journal of Cognitive Neuroscience, 21(1)*, 169-179. doi: 10.1162/jocn.2009.21014

Patterson, A., Seidenberg, M., & McClelland, J. (1989). Connections and Disconnections: Acquired Dyslexia in Computational Model of Reading Processes. In R. Morris (Ed.), *Parallel Distributed Processing: Implications for Psychology and Neurobiology* (pp. 131-181). Oxford: Clarendon Press.

Peerman, R., Dufour, S., & Burt, J. (2009). Orthographic Influences in Spoken-Word Recognition: The Consisteny Effect in Semantic and Gender Categorizaton Tasks. *Psychonomic Bulletin & Review, 16*, 363-368. doi: 10.3758/PBR.16.2.363

Peirce, J. (2002-2018). *Psychopy [Software]*. Retrieved at 2020, Dec. 16 from Psychopy: http://www.psychopy.org

Peirce, J., Gray, J., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., . . . Lindeløv, J. (2019). PsychoPy2: Experiments in Behavior Made Easy. *Behav. Res. Methods, 51(1)*, 195-203. doi:10.3758/s13428-018-01193-y

Perfetti, C. (1997). The Psycholinguistics of Spelling and Reading. In C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to Spell: Research, Theory, and Practice Across Languages* (pp. 21-38). Lawrence Erlbaum Associates.

Perfetti, C., & Hart, L. (2002). The Lexical Quality Hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of Functional Literacy* (pp. 189-213). Amsterdam: John Benjamins Publishing Company.

Perfetti, C., & Zhang, S. (1995). Very Early Phonological Activation in Chinese Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition 21*, 24-33. doi: 10.1037/0278-7393.21.1.24

Perfetti, C., Beck, I., Bell, L., & Hughes, C. (1987). Phonemic Knowledge and Learning to Read Are Reciprocal: A Longitudinal Study of First Grade Children. *Merrill-Palmer Quarterly, 33(3)*, 283-319. https://www.jstor.org/stable/23086537

Perfetti, C., Van Dyke, J., & Hart, L. (2001). The Psycholinguistics of Basic Literacy. *Annual Review of Applied Linguistics, 21*, 127-149. doi:10.1017/S0267190501000083

Perfetti, C., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a "Universal" Phonological Principle. In L. Katz & R. Frost (Eds.), *Orthography, Phonology, Morphology, and Meaning* (pp. 227-248). Amsterdam: North-Holland.

Perre, L., & Ziegler, J. (2007). On-line Activation of Orthography in Spoken Word Recognition. *Brain Research, 188,* 132-138. doi: 10.1016/j.brainres.2007.10.084

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. (2009). Orthographic Effects in Spoken Language: On-line Activation or Phonological Restructuring? *Brain Research 1275*, 73-80. doi: 10.1016/j.brainres.2009.04.018

Petersson, K., Ingvar, M., & Reis, A. (2009). Language and Literacy from a Cognitive Neuroscience Perspective. In D. Olson & N. Torrance (Eds.), *The Cambridge Handbook of Literacy* (pp. 152-182). Cambridge, UK: Cambridge University Press.

Petersson, K., Reis, A., & Ingvar, M. (2001). Cognitive Processing in Literate and Illiterate Subjects: A Review of Some Recent Behavioral and Functional Neuroimaging Data. *Scand. J. Psychol., 42*, 251-267. doi: 10.1111/1467-9450.00235

Petersson, K., Reis, A., Askelöf, S., Castro-Caldas, A., & Ingvar, M. (2000). Language Processing Modulated by Literacy: A Network Analysis of Verbal Repetition in Literate and Illiterate Subjects. *Journal of Cognitive Neuroscience, 12(3)*, 364-382. doi:10.1162/089892900562147

Pisoni, D., Nusbaum, H., Luce, P., & Slowiaczek, L. (1985). Speech Perception, Word Recognition and the Structure of the Lexicon. *Speech Communication, 4*, 75-95. doi: 10.1016/0167-6393(85)90037-8.

Pitt, M., & Samuel, A. (1990). The Use of Rhythm in Attending to Speech. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 564-573. doi: 10.1037//0096-1523.16.3.564

Plaut, D., McClelland, J., Seidenberg, M., & Patterson, K. (1996). Understanding Normal and Impaired Word Reading: Computational Principles in Quasi-Regular Domains. *Psychological Review, 103(1)*, 56-115. doi: 10.1037/0033-295x.103.1.56

Posner, M., & Snyder, C. (1975). Attention and Cognitive Control. In R. Solso (Ed.), *Information Processing and Cognition. The Loyola Symposium* (pp. 55-86). Lawrence Erlbaum Associates, Inc.

Prakash, P., Rekha, D., Nigam, R., & Karanth, P. (1993). Phonological Awareness, Orthography and Literacy. In R. Scholes (Ed.), *Linguistic and Cognitive Perspectives* (pp. 55-70). Lawrence Erlbaum Associates, Inc.

R Core Team. (2014). *R: A Language & Environment for Statistical Computing [software]*. Vienna, Austria. Retreived from R Foundation for Statistical Computing: https://www.R-projects.org

Rahbari, N. (2019). Error Patterns in Word Reading and Spelling in Persian: Is Spelling More Difficult than Reading? *Reading and Writing, 32*, 585-601. doi: 10.1007/s11145-008-9122-1

Rahbari, N., & Senechal, M. (2008). Lexical and Non-lexical Processes in the Skilled Reading and Spelling of Persian. *Springer Science, 22*, 511-529. doi: 10.1007/s11145-008-9122-1

Rahbari, N., & Senechal, M. (2010). Learning to Read and Spell in Persian: A Cross-sectional Study from Grades 1 to 4. *Developmental Psychology, 46(6)*, 1514–1527. doi: 10.1037/a0020377

Rahmani, H., Bijankhan, M., & Peyvasteh, S. (2007). Tosif-e Faza-ye Dark-e Vake Dar Zaban-e Farsi-ye Me'yar (A Investigation on the Perceptual Space for Vowels in Standard Persian). *Pazhuheshha-ye zabani (Linguistic Studies), 2*, 39-58.

Rapp, B., Benzing, L., & Caramazza, A. (1997). The Autonomy of Lexical Orthography. *Cognitive Neuropsychology, 14*, 71-104. doi: 10.1080/026432997381628

Rastle, K. (2012). Visual Word Recognition. In G. Hickok & S.L. Small (Eds.), *Neurobiology of Language* (pp. 255-264) . Academic Press. doi: 10.1093/oxfordhb/9780198568971.013.0005.

Rastle, K., & Brysbaert, M. (2006). Masked Phonological Priming Effects in English: Are They Real? Do They Matter? *Cognitive Psychology, 53*, 97-145. doi: 10.1016/j.cogpsych.2006.01.002

Rastle, K., McCormick, S., Bayliss, L., & Davis, C. (2011). Orthography Influences the Perception and Production of Speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(6)*, 1-22. doi: 10.1037/a0024833

Reis, A., & Castro-Caldas, A. (1997). Illiteracy: A Bias for Cognitive Development. *Journal of the International Neuropsychological Society, 3(5)*, 444-450. doi: 10.1017/S135561779700444X

Reis, A., Guerreiro, M., & Castro-Caldas, A. (1994). Influence of Educational Level of Non Brain-Damaged Subjects on Visual Naming Capacities. *Journal of Clinical and Experimental Neuropsychology, 16(6)*, 939-942. doi: 10.1080/01688639408402705

Roelofs, A. (2006). The Influence of Spelling on Phonological Encoding in Word Reading, Object Naming, and Word Generation. *Psychonomic Bulletin and Review, 13*, 33-37. doi: 10.3758/BF03193809

Rohani-Rahbar, E. (2012). Aspects of Persian Phonology and Morpho-Phonology. *Thesis*. Torento: University of Torento.

Rosenthal, J., & Ehri, L. (2008). The Mnemonic Value of Orthography for Vocabulary Learning. *Journal of Educational Psychology, 100*, 175-191. doi: 10.1037/0022-0663.100.1.175

Rosner, J., & Simon, D. (1971). The Auditory Analysis Test: An Initial Report. *Journal of Learning Disabilities, 4*(7), 384–392. doi: 10.1177/002221947100400706

Roux, S., & Bonin, P. (2013). With a Little Help from My Friends: Orthographic Influenes in Spoken Word Recognition. *Topics in Cogntive Psychology*, *113*, 35-48. doi: 10.3917/anpsy.131.0035

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for Phonemic Recoding in Visual Word Recognition. *Journal of Verbal Learning & Verbal Behavior, 10*(6), 645–657. doi:10.1016/S0022-5371(71)80071-3

Russell, V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, 69*, 1-29. doi: 10.18637/jss.v069.i01

Saletta, M. (2019). Orthography and Speech Production in Children with Good or Poor Reading Skills. *Applied Psycholinguistics, 40*, 905-931. doi:10.1017/S0142716419000055

Saletta, M., Goffman, L., & Hogan, T. (2016). Orthography and Modality Influence Speech Production in Adults and Children. *Journal of Speech, Language, and Hearing Research, 59(6),* 1421-1435. doi: 10.1044/2016_JSLHR-L-15-0242

Saletta, M., Goffman, L., & Hogan, T. P. (2016). Orthography and Modality Influence Speech Production in Adults and Children. *Journal of Experimental Psychology. Learning Memory and Cognition 37(6):* 1588-1594. doi: 10.1044/2016_JSLHR-L-15-0242

Samareh, Y. (1977). *The Arrangment of Segmental Phonemes in Farsi.* Tehran, Iran: University of Tehran Press.

Samareh, Y. (1985). *Avashenasi-e Zaban-e Farsi (Persian phonology).* Tehran, Iran: Markaz-e Nashr-e Daneshgahi (Center of University Press).

Schneider, K., Dogil, G., & Möbius, B. (2011). Reaction Time and Decision Difficulty in the Perception of Intonation. In *Twelfth Annual Conference of the International Speech Communication Association,* Florance, Italy, 27-31 August, 2011. http://www.isca-speech.org/archive/interspeech_2011/i11_2221.html

Scliar-Cabral, L., Morais, J., Nepomuceno, L., & Kolinsky, R. (1997). The Awareness of Phonemes: So Close-So Far Away. *International Journal of Psycholinguistics, 13(3)*, 211–240.

Sebastian, E., & Maldonado, A. (1986). Eldesarrollo de la capacidad de segmentacion de palabaras en lectores jovenes (The Development of Word Segmentation Ability in Early Readers). In J. Meisel (Ed.), *Adquisicion del lenguaje- Adequisicao da linguagem.* Frankfurt, Main: Dieter Verlag.

Seidenberg, M., & Tanenhaus, M. (1979). Orthographic Effects on Rhyme Monitoring. *Journal of Experimental Psychology, Human Learning, and Memory, 5*, 546-554. doi: 10.1037/0278-7393.5.6.546

Seidenberg, M., Waters, G., & Barnes, M. (1984). When Does Irregular Spelling or Pronunciation Influence Word Reading? *Journal of Verbal Learning and Verbal Behavior, 23*, 383-404. doi: 10.1016/S0022-5371(84)90270-6

Seymour, P., Aro, M., & Erskine, J. (2003). Foundation Literacy Acquisition in European Orthographies. *British Journal of Psychology, 94*: 143-174. doi: 10.1348/000712603321661859

Sheikh Sangtajan, S., & Bijankhan, M. (2013). Barrasi-ye Mahiyat-e Soutshenakhti-ye Vake-ha-ye Farsi dar Heja-ha-ye Goftar-e Ezhari (A Study on the Acoustic Nature of Persian Vowels in Spoken syllables). *Zabanpazhuhi (The Investigation of Language), 8*, 97-116.

Shields, J., McHugh, A., & Martin, J. (1974). Reaction Time to Phoneme Targets as a Function of Rhythmic Cues in Continuous Speech. *Journal of Experimental Psychology, 102*(2), 250–255. doi: 10.1037/h0035855

Shimron, J. (2015). *Reading Hebrew: The Language and the Psychology of Reading It.* New York: Routledge.

Shimron, J., & Navon, D. (1982). The Dependence on Graphemes and on Their Translation to Phonemes in Reading: A Developmental Perspective. *Reading Reasearh Quarterly, 17(2)*, 210-228. doi: 10.2307/747484

Siew, C., & Vitevitch, M. (2015). Spoken Word Recognition and Serial Recall of Words From Components in the Phonological Network. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(3),* 394–410. doi: 10.1037/xlm0000139

Siew, C., & Vitevitch, M. (2019). The Phonographic Language Network: Using Network Science to Investigate the Phonological and Orthographic Similarity Structure of Language. *American Phonological Association, 148*, 475-500. doi:10.1037/xge0000575

Silva, C., Faisca, L., Ingvar, M., Petersson, K., & Reis, A. (2012). Literacy: Exploring Working Memory Systems. *34*, 369-377. doi: 10.1080/13803395.2011.645017

Simpson, G., & Burgess, C. (1985). Activation and Selection Processes in the Recognition of Ambiguous Words. *Journal or Experimental Psychology: Human Perception and Performance, 11(1)*, 28-39. doi: 10.1037/0096-1523.11.1.28

Slowiaczek, L., Nusbaum, H., & Pisoni, D. (1985). Phonological Priming in Auditory Word Recognition. *Journal of Experimental Psychology Learning Memory and Cognition, 13(1),* 64-75. doi: 10.1037/0278-7393.13.1.64

Slowiaczek, L., Soltano, E., Wieting, S., & Bishop, K. (2003). An Investigation of Phonology and Orthography in Spoken-Word Recognition. *The Quarterly Journal of Experimental Psychology*, *56(2)*, 233-262. doi: 10.1080/02724980244000323

Smith, L., & Kemler, D. (1977). Developmental Trends in Free Classification: Evidence for New Conceptualization of Perceptual Development. *Journal of Experimental Child Psychology, 24*, 279-298. doi: 10.1016/0022-0965(77)90007-8

Smith, R. (2015). Perception of Speaker-Specific Phonetic Details. In S. Fuchs, D. Pape, C. Petrone, & P. Perrier (Eds.), *Individual Differences in Speech Production and Perception* (pp.11-38). Frankfurt am Main: Peter Lang.

Spence, C. (2008). Multisensory Perception. *Sensation, Perception, and Attention, 2,* doi: 10.1002/9781119170174.epcn214

Spence, C., & Bayne, T. (2015). Is Consciousness Multisensory? In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and Its Modalities* (pp. 95–132). Oxford University Press.

Spence, C., & Deroy, O. (2013). Crossmodal Mental Imagery. In Lacey, S., Lawson, R. (Eds.) Multisensory Imagery. NY: Springer. doi:10.1007/978-1-4614-5879-1_9

Stanovich, K., & Bauer, D. (1978). Experiments on the Spelling-to-Sound Regularity Effect in Word Recognition. *Memory and Cognition, 6*, 410-415. doi: 10.3758/BF03197473

Steinbrink, C., Groth, K., Lachmann, T., & Riecker, A. (2012). Neural Correlates of Temporal Auditory Processing in Developmental Dyslexia during German Vowel Length Discrimination: An fMRI Study. *Brain and Language, 121*, 1-11. doi: 10.1016/j.bandl.2011.12.003

Steyvers, M., & Tenenbaum, J. (2005). The Large-scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science, 29(1)*, 41–78. doi: 10.1207/s15516709cog2901_3

Stone, G., Vanhoy, M., & Van Orden, G. (1997). Perception Is a Two-way Street: Feedforward and Feedback Phonology in Visual Word Recognition. *Journal of Memory and Language, 36*, 337-359. doi: 10.1006/jmla.1996.2487

Strogatz, S. (2001). Exploring Complex Networks. *Nature, 410*, 268–276. doi: 10.1038/35065725

Swinney, D., & Prather, P. (1980). Phonemic Identification in a Phoneme Monitoring Experiment: The Variable Role of Uncertainty about Vowel Contexts. *Perception & Psychophysics*, *27*, 104–110. doi: 10.3758/bf03204296

Taft, M. (1992). The Body of BOSS: Subsyllabic Units in the Lexical Processing of Polysyllabic Words. *Journal of Experimental Psychology, Human Perception, and Performance, 18*, 1004-1014. doi: 10.1037/0096-1523.18.4.1004

Taft, M. (2011). Orthographic Influences When Processing Spoken Pseudowords: Theoretical Implications. *Frontiers in Psychology, 2*, 1-7. doi:10.3389/fpsyg.2011.00140

Taft, M., & Hambly, G. (1985). The Influence of Orthography on Phonological Representations in the Lexicon. *Journal of Memory and Language, 24*, 320-335. doi: 10.1016/0749-596X(85)90031-2

Taft, M., & Hambly, G. (1985). The Influence of Orthography on Phonological Representations in the Lexicon. *Journal of Memory and Language, 24*, 320-335. doi:10.1016/0749-596X(85)90031-2

Taft, M., Castles, A., Davis, C., Lazendic, G., & Ngye-Hoan, M. (2008). Automatic Activation of Orthography in Spoken Word Recognition: Pseudohomograph Priming. *Journal of Memory and Language, 58(2)*, 366-379. doi: 10.1016/j.jml.2007.11.002

Taha, H., & Azaizah-Seh, H. (2017). Visual Word Recognition and Vowelization in Arabic: New Evidnce from Lexical Decision Task Performances. *Cognitive Process, 18(4)*, 521-527. doi: 10.1007/s10339-017-0830-9

Toosarvandani, M. (2004). Vowel Length in Modern Farsi. J*ournal of the Royal Asiatic Society, 14*, 241-251. http://www.jstor.org/stable/25188473

Treiman, R. (1983). The Structure of Spoken Syllables: Evidence from Novel Word Games. *Cognition, 15*, 49-74. doi: 10.1016/0010-0277(83)90033-1

Treiman, R. (1991). Phonological Awareness and Its Role in Learning to Read and Spell. In D. Sawyer & B. Fox (Eds.), *Phonological Awareness in Reading. The Evolution of current Perspective* (pp. 159-189). New York: Springer-Verlag.

Treiman, R. (1993). *Beginning to Spell: a Study of First-grade Children.* Oxford: Oxford University Press.

Treiman, R., & Baron, J. (1981). Segmental Analysis Ability: Development and Relation to Reading Ability. In G. MacKinnon & T. Waller (Eds.), *Reading Research: Advances in Theory and Practice* (Vol. 3, pp. 159-198). New York: Academic Press.

Treiman, R., & Cassar, M. (1997). Can Children and Adults Focus on Sound as Opposed to Spelling in a Phoneme Counting Task. *Developmental Psychology, 33(5)*, 771-780. doi: 10.1037//0012-1649.33.5.771

Treiman, R., & Zukowsky, A. (1991). Levels of Phonological Awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological Processes in Literacy. A Tribute to Isabelle Y. Liberman* (pp. 67-83). Lawrence Erlbaum Associates, Inc.

Tsukada, K. (2011). The Perception of Arabic and Japanese Short and Long Vowels by Native Speakers of Arabic, Japanese, and Persian. *Journal of the Acoustical Society of America, 129*, 989-998. doi:10.1121/1.3531801

Tunmer, W., & Nesdale, A. (1985). Phonemic Segmentation Skill and Beginning Reading. *Journal of Educational Psychology, 77(4)*, 417-527. doi: 10.1037/0022-0663.77.4.417

Tunmer, W., & Rohl, M. (1991). Phonological Awareness and Reading Acquisition. In D. Sawyer & B. Fox (Eds.), *Phonological Awareness in Reading. The Evolution of Current Perspective* (pp. 1-30). New York: Springer-Verlag.

Van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of Letters and Speech Sounds in the Human Brain. *Neuron, 43*, 271–282. doi: 10.1016/j.neuron.2004.06.025

Van Ooijen, B., Cutler, A., & Norris, D. (1992). Detection of Vowels and Consonants with Minimal Acoustic Variations. *Speech Communication, 11*, 101-108. doi: 10.1016/0167-6393(92)90004-Q

Van Orden, G. (1987). A Rows Is a Rose: Spelling, Sound and Reading. *Memory and Cognition, 15*, 181-198. doi: 10.3758/BF03197716

Van Orden, G. (1990). Phonologic Mediation is Fundamental to Reading. In D. Besner & G. Humphreys (Eds.), *Basic Processes in Reading* (pp. 77-103). Lawrence Erlbaum Associates, Inc.

Van Orden, G., & Goldinger, S. (1994). The Interdependence of Form and Function in Cognitive Systems Explains Perception of Printed Words. *Journal of Experimental Psychology, Human Perception, and Performance*, *20*, 1269-1291. doi: 10.1037/0096-1523.20.6.1269

Van Orden, G., Jansen op de Haar, M., & Bosman, A. (1997). Complex Dynamic Systems Also Predict Dissociations, but They Do not Reduce to Autonomous Components. *Cognitive Neuropsychology, 14*, 131-165. doi: 10.1080/026432997381646

Vanhoy, M., & Van Orden, G. C. (2001). Pseudohomophones and Word Recognition. *Mamory & Cognition*, *29,* 522–529. doi:10.3758/BF03196403

Ventura, P., Kolinsky, R., Brito-Mendes, C., & Morais, J. (2001). Mental Representations of the Syllable Internal Structure Are Influenced by Orthography. *Language & Cognitive Processes, 16(4)*, 393-418. doi: 10.1080/01690960042000184

Ventura, P., Kolinsky, R., Pattamadilok, C., & Morais, J. (2008). The Developmental Turnpoint of Orthographic Consistency Effect in Speech Recognition. *Journal of Experimental Child Psychology, 100*, 135–145. doi:10.1016/j.jecp.2008.01.003

Ventura, P., Morais, J., & Kolinsky, R. (2007). The Development of the Orthographic Consistency Effect in Speech Recognition: From Sublexical to Lexical Involvement. *Cognition, 105(3)*, 547-576. doi:10.1016/j.cognition.2006.12.005

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The Locus of the Orthographic Consistency Effect in Auditory Word Recognition. *Language, Cognition and Neuroscience, 19(1),* 57-95. doi: 10.1080/01690960344000134

Vitevitch, M. (2008). What Can Graph Theory Tell Us about Word Learning and Lexical Retrieval? *Journal of Speech, Language, and Hearing Research, 51*, 408–422. doi: 10.1044/1092-4388(2008/030

Vitevitch, M., Goldstein, R., Siew, C., & Castro, N. (2014). Using Complex Networks to Understand the Mental Lexicon. *Yearbook of the Poznań Linguistic Meeting*, *1(1)*, 119-138 doi:0.1515/yplm-2015-0007

Wagner, R., Torgesen, J., Rashotte, C., & Pearson, N. (1999). *Comprehensive Test of Phonological Processing.* Austin, Texas: Pro-Ed.

Walley, A. (1993). The Role of Vocabulary Development in Children's Spoken Word Recognition and Segmentation Ability. *Developmental Review, 13*, 286–350. doi: 10.1006/drev.1993.1015

Walley, A., & Flege, J. (2000). Effects of Lexical Status on Native and Nonnative Vowel Perception: A Developmental Study. *27*, 307-332. doi: 10.1006/jpho.1999.0098

Watson, B., & Miller, T. (1993). Auditory Perception, Phonological Processing, and Reading Ability/ Disability. *Journal of Speech and Hearing Research, 36*, 850-863. doi: 10.1044/jshr.3604.850

Watson, J. (1919). *Psychology from a Standpoint of a Behaviorist.* Philadelphia: J.P. Lippincott.

Watt, T. (21. June 1954). *Brush up Your English*. Retrieved at 2022, Jan. 23 from https://linguistlist.org/issues/13/13-3353.html

Weber, A., & Scharenborg, O. (2012). Models of Spoken Word Recognition. *Cognitive Science, 3*, 387-401. doi: 10.1002/wcs.1178

Wimmer, H., Landerl, K., & Schneider, W. (1994). The Role of Rime Awareness in Learning to Read a Regular Orthography. *British Journal of Developmental Psychology, 12(4)*, 469–484. doi: 10.1111/j.2044-835X.1994.tb00648.x

Windfuhr, G. (1979). *Persian Grammar: History and State of Its Study.* Trends in Linguistic. State-of-the-Art Reports, 12. Mouton Publishers.

Windfuhr, G. (1990). Persian. In B. Comrie (Ed.), *The Major Languages of south Asia, the Middle East, and Africa* (Chapter 6), London: Routledge. doi: https://doi.org/10.4324/9780203412336

Xu, B., & Perfetti, C. A. (1999). Nonstrategic Subjective Threshold Effects in Phonemic Masking. *Memoory & Cognition, 27(1)*, 26-36. doi:10.3758/BF03201210

Ziegler, J., & Ferrand, L. (1998). Orthography Shapes the Perception of Speech: The Consistency Effect in Auditory Word Recognition. *Psychonomic Bulletin & Review, 5*, 683-689. doi:10.3758/BF03208845

Ziegler, J., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin, 32*, 3-29. doi:10.1037/0033-2909.131.1.3

Ziegler, J., Ferrand, L., & Montant, M. (2004). Visual Phonology: The Effect of Orthographic Consistency on Different Auditory Word Recognition Tasks. *Memory & Cognition, 32*, 732-741. doi:10.3758/BF03195863

Ziegler, J., Muneaux, M., & Grainger, J. (2003). Neighborhood Effects in Auditory Word Recognition: Phonological Competition and Orthographic Facilitation. *Journal of Memory and Language, 48(4),* 779-793. doi:10.1016/S0749-596X(03)00006-8

Ziegler, J., Petrova, A., & Ferrand, L. (2008). Feedback Consistency Effects in Visual and Auditory Word Recognition: Where Do We Stand After More than a Decade? *Journal of Experimental Psychology Learning Memory and Cognition, 34*, 643-661. doi: 10.1037/0278-7393.34.3.643

Curriculum Vitae

Leila Nayernia

Job experience

2002-2010	Instructor in Linguistics at the University of Applied Sciences, Shiraz, Iran.
2002-2019	Linguist at Future Group.
2017-2018	Linguist at LingoKing.
2020-2022 (March)	Scientific staff and instructor in Linguistics at Heinrich Heine University, Düsseldorf, Germany.

Education

1994-1997	Bachelor student in Translation. S & B University, Zahedan, Iran.
1998-2002	Master student in Philology, Iran, Shiraz. Thesis topic: Syntactic structures of Persian adjectives.
2012-2017	Master student in Linguistics, Heinrich Heine University, Düsseldorf, Germany. Thesis topic: The influence of orthography on phonological processing.
2017-date	Ph.D. candidate in linguistics, focusing on Psycholinguistics, Heinrich Heine University, Düsseldorf, Germany. Research topic: The influence of orthography on phonological processing.

Eidesstattliche Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der "Ordnung über die Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist.

Düsseldorf, 09.09.2022

(Ort, Datum) (Unterschrift)