Opening Your Ears: The Role of L1 in Processing of Nonnative Prosodic Contrasts

Anna Lukyanchenko, William J. Idsardi, and Nan Jiang
University of Maryland

1. Introduction

A wealth of well-documented evidence in the SLA literature demonstrates that the way second language (L2) learners perceive and process foreign speech is mediated by the properties of their native language (L1). By far, the majority of inquiries have focused on the problem of acquisition of L2 segmental features by L2 learners. However, languages differ not only in their repertoire of segmental elements (vowels and consonants), but also in their repertoire and use of suprasegmental properties (these are not necessarily confined to or discretely ordered with any particular segment, but occur at some higher levels in an utterance, like syllable, word, phrase, etc.). That is why mastering L2 phonology includes not only acquisition of L2 segments, but also acquisition of L2 prosody and learning how to integrate phonetic and prosodic information in a native-like manner. By investigating the role a speaker's first language plays in the processing of nonnative prosodic contrasts as far as primary word stress is concerned, the present study seeks to shed more light on the mechanism underlying the formation of abstract L2 prosodic representations, and to explain how this process might be *grammatically* or *perceptually* constrained or facilitated depending on the speaker's L1. Specifically, the processing of contrastive word stress by speakers of three typologically different languages is examined: French, Russian, and Persian.

The paper is organized as follows: Section 2 presents an overview of the theoretical and empirical literature and outlines two major challenges learners face in the acquisition and processing of word stress: the grammar learning problem and the perception problem. Section 3 provides a brief description of word stress in the three languages. Section 4 describes the tasks used in the current study, and in Section 5, the results are discussed in light of the Parameter setting theory (Chomsky, 1981; Archibald, 1993) and the current psycholinguistic approach to stress perception (Dupoux et al., 1997; 2001; 2008.). Finally, Section 6 concludes the paper.

2. Learning L2 word stress

In the present study, word stress refers to the perceived prominence of a syllable in a polysyllabic word, and the way it manifests itself in speech is highly language-dependent (both in terms of its placement in a word and the acoustic cues used to realize it). For example, there are languages like English and Russian, which have varying patterns of stress assignment, and languages like Finnish and Hungarian, where stress falls on the first syllable, or Polish, where stress is usually on the penultimate syllable. Similarly, the use of different acoustic cues to make one designated syllable in a word most prominent also varies from language to language. There are languages that use dynamic accent cues (e.g., pitch, loudness, intensity), qualitative prosodic cues (e.g., change of vowel quality), quantitative cues (e.g., duration, vowel reduction), or a combination of those[1].

---

* We are grateful to the three anonymous reviewers of the abstract, two anonymous reviewers of the manuscript and the SLRF 2010 audience for their insightful comments and suggestions.

[1] In the context of metrical theory, there are no absolute values that can be assigned to stressed or unstressed syllables; such values are identified based on the prominence relationships among constituents (Liberman, 1975; Liberman & Prince, 1977).

Taking into account such variability concerning realization of word stress across languages, the important question that one might reasonably ask, then, is how L2 learners acquire nonnative stress patterns. When children first learn a language, they learn to extract rhythmical-periodical properties of speech signal through the processes of phonetic and phonological encoding, and then learn how to map the surface form onto a given template/representation stored in the long-term memory (i.e., a metrical template). This facilitates language acquisition and triggers adjustments on perception and production routines that allow efficient language-specific processing (Cutler & Mehler, 1993; Speer & Ito, 2009).

In adult L2 acquisition, the perception and production routines are no longer very flexible because adult learners already have a prosodic system in place for their L1, so they have to discover how L2 system is consistent or inconsistent with the L1 system, and make necessary adjustments in their grammars. However, this is not a trivial task, and very often the differences between L1 and L2 manifest themselves in word stress errors in L2, thereby contributing to a detectable degree of foreign accent. For instance, L2 stress location may pose certain learning and processing difficulties when L1 stress location is different and may result in incorrect stress placement in L2 (Swan & Smith, 2001; Van der Pas & Zonneveld, 2004; Dupoux et al., 1997; 2001; 2008).

Constrained by the prosodic properties of their L1, learners also very often fail to make use of relevant acoustic cues in the production and perception of L2 stress patterns. For example, in a series of studies by Nguyen et al. (2005; 2008) it was found that Vietnamese learners can differentiate between stressed and unstressed syllables in English by means of F_0 contrast - an acoustic correlate available in both languages, but they fail to produce a syllable duration contrast that characterizes native productions and do not reduce vowels in unstressed syllables - possibly because these two important phonetic features are not active in Vietnamese tonal contrasts.

Similar empirical and anecdotal evidence makes L1 transfer errors widely acknowledged as an important subject of inquiry into the acquisition of L2 prosody in general, but we still need to be concerned with the question of *when* and *why* learners transfer *what* and *how much* (Ringbom, 1990). In the current study, we propose that transfer errors might have two sources of origin: the inability to reset relevant parameters and the perception problem. It is reasonable to suppose that if the learner is not able to perceive L2 stress correctly in the first place, the input will not act as a triggering factor, and the relevant parameters will not be reset. In order to tease these two problems apart and try to understand learning and processing difficulties in L2, examination of typologically different languages with different metrical parameters might prove to be useful.

2.1. Grammar learning problem

Within a Universal Grammar (UG) approach, the learner is viewed as having a set of certain innate universal principles unaffected by the linguistic environment and a set of parameters which need to be set depending on the language the learner is exposed to. Numerous studies have shown that a principles-and-parameters approach is well suited to the study of word stress because it provides a systematic account of the possible stress patterns in the world’s languages (e.g., Liberman & Prince, 1977; Halle & Vergnaud, 1987; Dresher & Kaye, 1990; Idsardi, 1992; Pater, 1997). As suggested by Dresher & Kaye (1990), most metrical (stress) parameters appear to be binary:

P1: Word tree	[Left/Right]
P2: Foot type	[Binary/Unbounded]
P3: Strong on	[Left/Right]
P4: Built from	[Left/Right]
P5: Quantity-sensitive	[Yes/No]
P6: Sensitive to	[Rime/Nucleus]
P8: Extrametrical	[No/Yes]
P8A: Extrametrical on	[Left/Right]

Given the complexity of the system with various parameters, correct stress assignment is dependent on setting all or most parameters correctly as opposed to just using a single, general stress rule (Dresher & Kaye, 1990). Such fact has direct implications for L2 acquisition. For example, if L1 and L2 have the same parameters, the parameter setting may proceed through a feature copying mechanism: $f(x) = x$. However, in the case when L1 and L2 do not share the same parameters, the learner will have to make certain changes in the mental representation of stress by resetting one or several L1 metrical parameters. In this situation, the learner has to learn the new rules, which would look somewhat like the following: If you find *x* where you were expecting *y*, change parameter *z* (Archibald, 1998).

With regard to the resetting of metrical parameters, there exist controversial and unanswered questions as to whether all parameters can be reset successfully in L2 acquisition, whether bilinguals can juggle contradicting L1 and L2 parameters, and, if yes, how one parametric value can take precedence over the other depending on the linguistic context. Relevant to this, Van der Pas and Zonneveld (2004) argue that in the case of L2 with parameters different from those of the speaker's native language, parameter resetting is impossible as far as the metrical system is concerned, especially in those cases in which acquisition of L2 - as it often does - takes place after puberty. Instead of successful parameter resetting, learners over-generalize their previously learned L1 rules and apply them to their L2.

A different tendency was observed by Pater (1997), who found that L1 French speakers learning L2 English seem to be "missetting" L2 parameters in the sense that their production of primary word stress in English suggests a parameter setting that is different from either L1 or L2.

Conciliatory evidence comes from numerous studies by Archibald (1993; 1995; 1997; 1998) who investigated the acquisition of L2 stress by speakers from different L1 backgrounds. According to him, adult L2 learners' interlanguages do not violate metrical universals and adults are capable of resetting their parameters to be consistent with the L2. He acknowledges, however, that L2 learners' grammars are fundamentally different from those of native speakers in a sense that their interlanguages are a combination of co-existent UG principles, correct L2 parameter settings (from parameter resetting), and incorrect L1 parameter settings (from L1 transfer). As a result, this might produce indistinguishable L2 behavior, but the underlying systems and representations for stress might still be different.

2.2. Perception problem

The investigation of the acquisition of L2 word stress has also been approached from a psycholinguistics and experimental psychology standpoint, which is mostly concerned with how L2 prosodic representations are built and used in speech processing. The previously mentioned characteristics of word stress such as its variable position in a word and the cross-linguistic variation in its acoustic correlates have spurred research investigating how L1 stress properties bear on the perception and production of L2 stress.

Several models have been proposed in this regard. The Stress Deafness Model (SDM) put forward by Peperkamp & Dupoux (2002) bases its predictions on the observed surface transparency (regularity) of L1 stress patterns and its possible implications for the mental representation of stress: the more regular the L1 stress system is (or the easier it is to observe that regularity), the more difficulties the native speakers of such language will have with the perception of stress (both in L1 and L2). The model further claims that a complete regularity of the L1 stress patterns leads to the loss of the phonological representation of stress and results in stress 'deafness'. According to this model, such stress 'deafness' occurs in perception at the level of phonological decoding and resists training (Dupoux et al., 1997).

In a similar vein, the Stress Typology Model (STM) proposed by Altmann and Vogel (2002) suggests that native speakers of stress languages with unpredictable stress experience the least problems with the perception of L2 stress, whereas speakers of languages with predictable stress have more difficulties with it. The underlying assumption is that regularity requires setting a larger number of L1 metrical parameters to positive values, which interferes with L2 stress perception (Altmann, 2006).

Kijak (2009) tested the predictions of the above models and found that the surface regularity versus irregularity of L1 stress patterns only partially accounts for stress perception ability, and that the function that stress performs in L1 (such as being relevant in speech segmentation or word recognition) may be crucial as well. For example, unpredictable and lexically encoded stress in Russian has an important contrastive function; therefore, Russian native speakers are good at stress perception. In English, stress only has a partially contrastive function and is not as crucial in word recognition, which has consequences for stress perception.

Summarizing the assumptions of the above models, dimensions that have a contrastive value as far as L1 stress is concerned (e.g., Russian), will be specified in full detail and represented at the abstract phonological level, whereas those that have highly predictable or fixed stress distribution (e.g., French) will be underrepresented or not represented at all. In this paper, we further explore the predictions made by the above models through examining stress perception behavior of speakers of three typologically different languages with varying degree of stress predictability in their L1s: French, Russian, and Persian.

3. Language groups

3.1. French

French is a Romance language with stress usually falling on the rightmost full vowel (Roca, 1999). However, some questions regarding French metrical phonology remain unresolved. Some theories treat French as a language with word final stress (Altmann, 2006; Pater, 1997; Van der Pas & Zonneveld, 2004), others categorize it as a phrasal stress, where the most prominent syllable is usually the last non-schwa syllable in the phonological phrase (which may or may not coincide with the word) (Dupoux et al., 2008; Tremblay & Owens, 2010; Roca, 1999; Kijak, 2009). It is typically realized as final syllable lengthening in prosodic groups with no significant increase in frequency or intensity. Thus, French stress does not play a contrastive function, but rather, has a demarcative role (Dupoux et al., 2008).

Based on the predictions of the outlined theoretical models, French speakers should have substantial difficulty in representing and acquiring contrastive stress, which has, in fact, been demonstrated by a series of studies by Dupoux et al. (1997; 2001; 2008; 2009) and Kijak (2009). It was shown that French participants can use the acoustic stress cues in order to perform standard discrimination tasks flawlessly, but nevertheless show ‘deafness’ to stress contrasts at a more abstract processing level when higher memory load and talker variability are added to the task. Thus, Kijak (2009) conducted a number of tasks investigating the acquisition of stress by L2 speakers of nine different languages and concluded that the French group was indeed the ‘worst’ group out of all tested languages as far as stress perception and production is concerned.

3.2. Russian

Russian belongs to the Slavic branch of Indo-European languages, and it is characterized by a free word stress, both in terms of its placement in the word and its relationship to morphemic structure (i.e., stress can fall on any syllable in the word and on any morpheme). Word stress in Russian serves a contrastive and a constitutive function, and is actively involved in word identification and recognition (Svetozarova, 1998).

Phonetically, Russian stress is characterized by a set of features, most important of which are the duration and the quality of the stressed vowel. Additionally, stressed vowels might also be associated with a higher intensity than unstressed vowels, which usually undergo significant qualitative and quantitative reduction.

Due to its ‘mobility’ and unpredictability, word stress in Russian can be described as a lexical accent system. It means that a significant majority of words in Russian receive stress through the process of mapping an inherent accent onto the surface form of a morpheme. The models we reviewed in the previous section predict that native speakers of languages whose stress system is almost completely lexical, and where stress is encoded at the abstract phonological level, should be very good at stress perception. Indeed, the results obtained by Kijak (2009), provide corroborating evidence: Russian speakers obtained the highest accuracy scores out of all tested language groups.

3.3. Persian

Studies on L2 stress perception up to date have been mostly restricted to a handful of languages, such as French, Polish, Hungarian among the languages with regular stress; Spanish and English among the languages with a less transparent stress system; and Chinese, Japanese, and Korean among the languages with no word-level stress in their L1. However, it is important that more languages are investigated on the matter of stress perception in order to see whether the findings can be generalized to speakers of other languages as well. To the best of our knowledge, no study has addressed the perception of word stress by Persian speakers, and the exact nature of the prominence attributed to word stress in Persian is unattested.

Persian is an Iranian language within the Indo-European family of languages and represents an interesting case for studying word stress. Predominantly, the last syllable of a word in Persian receives the stress. However, some degree of variability of the stress patterns in Persian has led linguists to suggest that word stress in Persian is not as regular as had been previously thought. The exceptions to the final-syllable stress rule concern certain lexical categories. Whereas in nouns and adjectives the main stress is on the final syllable, in verbs, the main stress sometimes occurs on the penultimate or on the initial syllable (Windfuhr 1990; Mahjani, 2003). Namely, exceptions to the final-syllable stress rule involve only words containing one or two of the four stressed verbal prefixes ˈmi-, ˈbe-, ˈnæ-, ˈmæ- (the first prefix is always stressed when there are several such prefixes in a word), and words containing one or more of certain unstressed suffixes (in this case, stress falls on the syllable immediately preceding the suffix) (Ferguson, 1957). The occurrence of such non-final verbal stress yields a fair number of minimal contrasts with noun forms, e.g., ˈbedehi (“(that) you give”) and bedeˈhi (“debt”), therefore, besides being inflectional, Persian stress can also play a contrastive role. Acoustically, syllable prominence is realized as the relative loudness between stressed and unstressed vowels. Additionally, vowel duration might provide an additional stress cue (Ferguson, 1957).

Table 1. An overview of the word stress in Russian, French, and Persian

	Russian	French	Persian
Type	lexical	phrasal	inflectional can be lexical
Acoustic/phonetic properties of word stress	duration vowel quality intensity vowel reduction	duration (*only final syllable lengthening in prosodic groups*)	intensity duration
Functional characteristics of word stress	contrastive constitutive	demarcative	morphological can be contrastive
Regularity	irregular unpredictable	regular predictable	quite regular predictable

Persian is a critical language for the purposes of this study because it constitutes a transitory group between French and Russian (see Table 1 for a comparative description of word stress in the three languages). By comparing these three languages, we want to see how the properties of the L1 metrical system affect processing of word stress patterns in a nonnative language, and what the consequences might be for the acquisition of contrastive word stress in L2.

Based on the findings from earlier studies, we hypothesize that Russian speakers would outperform speakers of French. We make no predictions, however, concerning processing of stress by Persian speakers. As has been demonstrated by Kijak (2009), stress ‘deafness’ is not limited to a single language (French), and can be extended to languages with predominantly regular stress patterns (e.g., Finnish). Thus, Persian speakers might also reveal perception difficulties and the inability to reset parameters similar to French and Finnish speakers due to the substantial regularity of the L1 stress system. Conversely, because there are some exceptions in Persian word stress system, Persian speakers might be able to establish phonological representation of word stress and have the necessary parameters in place to process nonnative stress with a relative ease.

4. Method

4.1. Participants

Nine Russian speakers, 10 Persian speakers, and 8 French speakers took part in the study. There were 6 female and 3 male participants in the Russian group (age 22-47), 4 female and 6 male participants in the Persian group (age 22-30), and 4 female and 4 male speakers of French (age 24-31). On average, the Russian speakers in the study have lived in the United States for 6.68 years, the Persian participants had a length of residence (LOR) in the U.S. for 4.75, and that of our French speakers was 2.63 years. All participants were advanced speakers of English. Most all of them reported high percentage of use of English in their daily lives and a small percentage of use of their L1s. All participants had to rate their L2 competence in several linguistic domains (pronunciation, grammar, speaking, listening, writing, reading) on a scale from 1 to 10. Average proficiency ratings per group are reported in Table 2. At the time of testing, the majority of participants were enrolled in graduate programs at the University of Maryland, College Park; some participants were also working in Washington, D.C. area. None of the subjects had a known hearing deficit.

Table 2. Participants' biographic information

L1	N	Mean age, yrs *(SD)*	LOR, yrs *(SD)*	Self-reported proficiency in L2 (1 is the lowest, 10 is the highest)
Russian	9	31.67 *(8.92)*	6.68 *(5.07)*	8.20
French	8	27 *(1.89)*	2.63 *(1.73)*	7.10
Persian	10	27.13 *(2.91)*	4.75 *(1.75)*	7.97

4.2. Materials

Stimulus materials for the two tasks used in the present study were identical. First, two disyllabic minimal pairs involving a stress contrast were constructed, with stress falling on the first syllable and stress falling on the second syllable. The first contrast, /ˈmipa – miˈpa/ (mean duration ~ 445 ms), contained two different vowels in the two syllables, and the other contrast, /ˈfiki – fiˈki/ (mean duration ~465 ms), had identical vowels in the two syllables. All items were phonotactically legal nonwords and contained only sounds present in the three languages[2].

Table 3. Mean differences between stressed and unstressed syllables

syllable	duration, ms		pitch, Hz		intensity, dB	
	stressed	unstressed	stressed	unstressed	stressed	unstressed
mi	97.25	61.75	179	153	67	62.5
pa	135.5	78.25	168.25	142	73	61.5
fi	93.75	60	201.25	146.5	66.25	59.5
ki	105.75	75.75	187.25	130.25	64.75	57.25

Second, stimuli were recorded by four native Spanish speakers (2 males, 2 females) who were trained to produce target stress patterns. Four talkers were used in this study in order to ensure that the

[2] There is a restriction on syllable structure in Persian: it does not allow vowels in the syllable onset position (even words that start with a vowel include the glottal stop as the syllable onset), so all stimuli had to be of the CVCV type.

acoustic level of processing is less readily available to the participants, and that the task is getting at the phonological level of representations. Spanish speakers were chosen because Spanish permits contrastive word stress. The major phonetic cue to stress in Spanish is vowel duration, but pitch and intensity might also play a role. Each token was recorded five times, and the recordings underwent acoustic analysis in PRAAT (Boersma & Weenink, 2010) so that only tokens with most unambiguous stress patterns were selected. For each talker, only one recording of each item was eventually chosen. Additionally, the word ''OK'' was recorded once by a female talker. All recordings were digitized at 44 kHz. On average, stressed vowels were significantly longer than unstressed vowels; they were also significantly louder and had a higher pitch[3] (see Table 3).

4.3. Task 1. Identification task
4.3.1. Design and procedure

The first task was a forced-choice identification task, and it consisted in identifying the placement of stress in the two types of nonwords, /mipa/ and /fiki/. Participants were tested individually on DELL computers in the Second Language Acquisition lab at the University of Maryland. DMDX software platform (programmed by K. Forster and J. Forster, the University of Arizona) was used to control the stimuli presentation and record response latencies to target words, timed from the onset of the target word. The experiment included two blocks: the MIPA block and the FIKI block. Participants were randomly assigned to either of the two groups so that half of the participants were first presented with the MIPA block and half of the participants were first exposed to the FIKI block. However, the later analysis revealed that there was no effect of the order of the presentation.

In each block, the testing phase was always preceded by the familiarization phase and procedural practice. At the familiarization phase, participants were told that they would hear two types of foreign words pronounced by different speakers and that their task consisted in learning to discriminate these words. First, they heard the nonword with stress on the first syllable (e.g., /ˈmipa/) and were told that this is Type1 word. They were given as much time as they needed to familiarize themselves with the nonword pronounced by the four talkers. The same procedure was repeated for the nonword having stress on the second syllable, i.e., /miˈpa/ (Type 2 word in the experiment). After this, participants were presented with both types of words: /ˈmipa – miˈpa/. The familiarization phase with both words continued for 1 minute or until participants informed the instructor that they were ready to proceed to the test. Before the testing phase, however, participants learned to associate number keys with responses: number key [1] for Type 1 word, number key [2] for Type 2 word. Participants were given 2 seconds to provide the response. If they did not press the button within 2 seconds, the presentation continued with the next item. Thus, they were instructed to respond as quickly but also as accurately as they could. The first eight items constituted procedural practice, and were not included in the analysis. The task consisted of 64 randomly presented nonwords (32 tokens for each type of the word). No feedback was provided at any phase of the task, and the whole procedure was repeated for each block (MIPA and FIKI). Overall, the entire experiment with two blocks lasted about 10-15 minutes.

4.3.2. Results

The data obtained from the identification task were analyzed along the two outcomes: reaction time and error rate. A mixed design ANOVA with L1 as a between subjects factor and Type of the word as a within subjects repeated factor was carried out[4]. The reaction time data analysis yielded a

---

[3] ANOVA results confirmed that most stressed vowels were significantly longer than unstressed vowels across all syllables: /mi/, $F(1,6) = 16.3$, $p=0.007$; /pa/, $F(1,6) = 14.35$, $p=0.009$; /fi/, $F(1,6) = 7.57$, $p=0.033$; except for /ki/, $F(1,6) = 4.79$, $p=0.07$. Stressed vowels were also louder than unstressed vowels: /pa/, $F(1,6) = 21.16$, $p=0.004$; /fi/, $F(1,6) = 23$, $p=0.003$; /ki/, $F(1,6) = 22.88$, $p=0.003$, except for /mi/, $F(1,6) = 3.32$, $p=0.11$. Because stimulus items were recorded by both male and female speakers, the differences in pitch between stressed and unstressed syllables were analyzed separately for male and female voices, and were found to be significant as well.

[4] Reaction time data and error rate data were also analyzed according to the Type of stress (the first vs. the second syllable), but the difference between the means was not statistically significant–participants responded to the

significant main effect for the L1 of the participants: F (2, 51) = 31.36, $p < 0.001$, but no significant main effect for the type of the word (MIPA or FIKI).

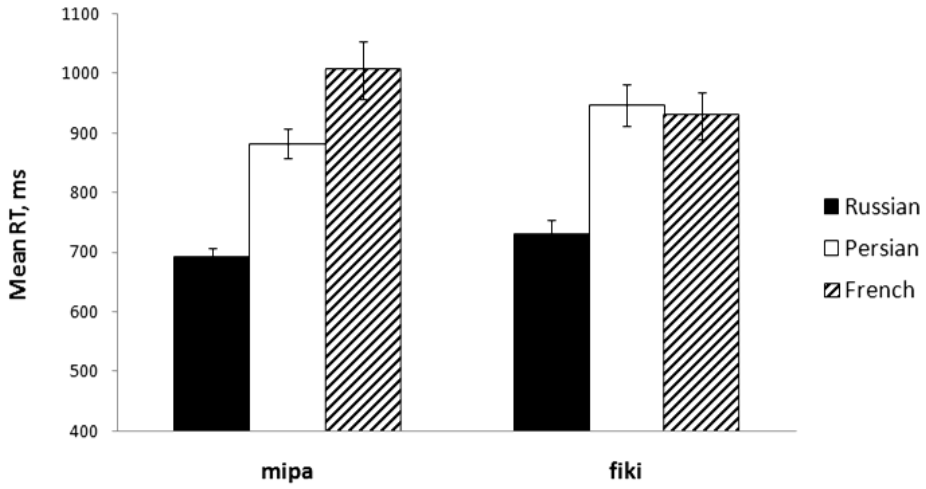

Figure 1. Identification task Mean Reaction Times (ms) across all language groups

The interaction between the two variables was not significant, which means that the main effect can be attributed to participants' L1 only. According to standard Cohen's (1988) criteria, the observed effect size was rather large ($\eta^2 = 0.552$), and observed power was maximal (1.0). In other words, Russian, French, and Persian speakers performed on the task differently, independent of whether they were discriminating between MIPA contrasts or FIKI contrasts (The results are summarized in Figure 1). In order to investigate this effect further, Dunn-Sidak post hoc comparisons were conducted. The analysis revealed that Russian participants were significantly faster than either Persian or French speakers ($p < 0.001$), but that performance of Persian and French participants did not differ significantly.

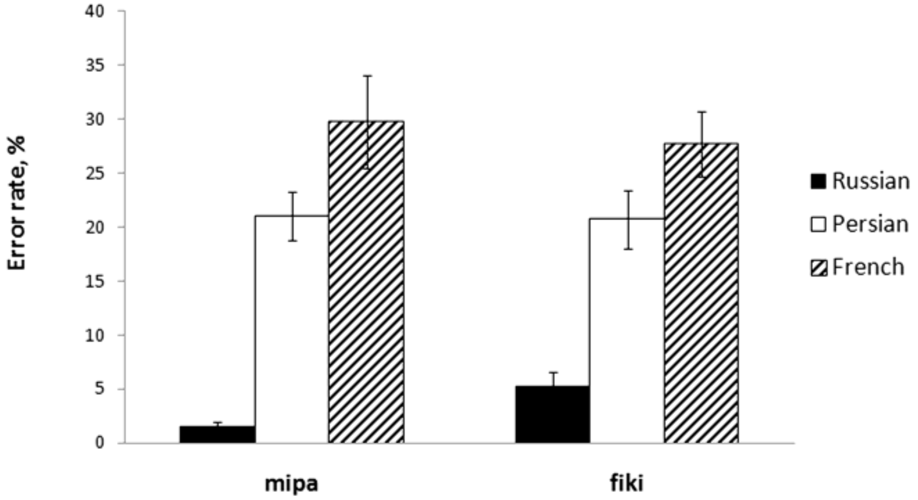

Figure 2. Identification task Mean Error Rate (%) across all language groups

---

nonwords with stress on the first syllable and to the nonwords with stress on the second syllable in a similar manner, so the data were combined for further analysis.

In the error rate data analysis, the ANOVA yielded a significant main effect for the L1 of the participants, $F(2, 51) = 30.49$, $p < 0.001$. Again, the main effect for the word type (MIPA vs. FIKI) and the interaction between the two factors were not statistically significant. The observed effect size for L1 was rather large ($\eta^2 = 0.545$), and observed power was maximal (1.0). A series of post hoc multiple comparison procedures were administered to investigate the nature of the main effect. The results demonstrated that Russian participants were significantly more accurate in the identification of contrastive stress patterns than either Persian or French speakers ($p < 0.001$). However, there was also a difference in the performance of Persian and French speakers, which was marginally significant.

Because the same results were obtained for the usual, the Geisser-Greenhouse, and the Huynh and Feldt F tests, we assumed that the sphericity assumption was upheld and that the data were sufficiently close to spherical. This allowed us to run the Poisson general linear model. The error rate data were fit to a GLM with a Poisson distribution in JMP 9.0 software. The results of the planned comparisons of the Persian group with the other two groups revealed that Persian participants were different from both Russian ($p < 0.001$) and from French participants ($p < 0.01$). The results are presented in Figure 2.

It should be mentioned that he three groups of participants in the present study differed in their L2 proficiency and length of residence (LOR) in the U.S.: Russian speakers lived in the U.S. for 6.68 years, Persian speakers for 4.75 years, and French speakers for 2.63 years. Their self-reported L2 proficiency was 8.2, 7.97, and 7.1, respectively. However, the results of the linear regression analyses (with the participants' reaction time data and error rate data as dependent variables and length of residence (LOR) and proficiency as predictor variables) indicate that the total variation in the data cannot be predicted by LOR ($r^2 = 0.092$, $F(1,25) = 2.5$, $p = 0.125$, for reaction time, and $r^2 = 0.13$, $F(1,25) = 3.73$, $p = 0.07$, for error rate) or proficiency score ($r^2 = 0.044$, $F(1,25) = 1.15$, $p = 0.3$, for reaction time, and $r^2 = 0.06$, $F(1,25) = 1.56$, $p = 0.22$, for error rate). Thus, the difference in the participants' performance on the task can be attributed to cross-language differences in prosody as opposed to differences in the amount of experience with and proficiency in English, a second language for all our participants.

To summarize, in accordance with our predictions, Russian participants outperformed French and Persian participants on the Identification task on both task measures: their responses were not only more accurate, but also significantly faster. The performance of French speakers was the poorest among the three groups – such evidence also seems to be in agreement with the previous findings reviewed earlier. The performance of Persian participants points to the fact that they might indeed fall into a transitory category between Russian speakers and French speakers. Although their reaction time data were not statistically different from those of the French speakers, their accuracy data show a statistical difference, which suggests that they might still be processing stress differently from our French subjects. In order to obtain further evidence on processing of word stress by the three groups, a sequence recall task was carried out.

4.4. Task 2. Sequence recall task

4.4.1. Design and procedure

This task was a modification of the speeded sequence recall task used by Dupoux et al. (2001) and Dupoux et al. (2008). SRT is a short-term memory task, and it is designed to make the acoustic level of representation less accessible and hence to highlight processing at the phonological level. The task consisted in reproducing sequences of sounds presented auditorily by recoding them into corresponding numbers.

The same participants who participated in the identification task took part in the sequence recall task, which was always administered after the identification task. The block order (for /mipa/ and /fiki/) was accordingly balanced among the participants. Moreover, the same sound recordings as in Task 1 were used in this task so that each word (/ˈmipa/, /miˈpa/, /ˈfiki/ and /fiˈki/) was instantiated by one of the four acoustically different tokens pronounced by four different talkers. Although by the time of SRT, the participants were already familiar with the sounds and already learned the associations between the word types and number keys, they were allowed to listen to the recordings again before each block in order to refresh them in their memory.

For each nonword, MIPA and FIKI, four-item sequences were constructed for each of the two blocks, e.g., ˈfiki – fiˈki – fiˈki - ˈfiki [1-2-2-1], etc. Out of the 16 possible combinations of four items, only eight sequences with a greater transition were selected (as in Experiment 1 in Dupoux et al., 2001). For example, a sequence like 1-2-1-1 has more variation because it contains one transition from 1 to 2 and another one from 2 to 1. In contrast, a combination like 2-1-1-1 has only one transition from 2 to 1, and should be easier to capture. The following eight sequence types were used in the experiment: 6 two-transition sequences (1-1-2-1, 1-2-1-1, 1-2-2-1, 2-1-1-2, 2-1-2-2, 2-2-1-2) and 2 one-transition sequences (2-1-1-1, 1-1-2-2). Each sequence was repeated 3 times yielding 24 sequences overall. Six sequences with 2-item variation were included for practice at the beginning of each block. Sequences were randomized so that the same talker was never presented adjacently, and no single token appeared more than once in a sequence. In order to diminish the likelihood of participants' mentally translating the sound into the associated numbers while listening to the sequences, the silent period between the items in a sequence was kept very short, 50 ms (as in Dupoux et al. 2008), and the ISI between the sequences was held constant at 1500 ms. In addition, the word "okay" was included after the presentation of each sequence. Participants were instructed to listen to the whole sequence and write down the correct coding on the answer sheet only after they have heard the word "okay". This was done in order to prevent participants from using echoic memory (Dupoux et al., 2001) and from starting to write before the whole sequence was presented. No feedback was provided during the test phase. On average, the entire task with two blocks lasted about 10-15 minutes.

4.4.2. Results

Raw responses from the speakers of the three language groups were obtained and transformed into participants' error rate data. Responses that were a 100% correct transcription of the presented sequence were coded as correct. When there was at least one error in any of the token or token position in the recorded sequence, the response was coded as incorrect. Figure 3 shows mean error rate in the three language groups.

As evidenced by the graph, Russian speakers made fewer errors than Persian and French speakers, whose performance revealed high error rate (far more than 50%). Such observation was confirmed statistically by carrying out a two-factor split-plot ANOVA with L1 as a between subjects factor and Type of the word as a within subjects repeated factor. In agreement with the identification task, the ANOVA results for the SRT task yielded a significant main effect for the L1 of the participants, $F(2, 24) = 18.67$, $p < 0.001$, and no significant main effect for the word type and no significant interaction between the two variables. The observed effect size for L1 was very large ($\eta^2 = 0.609$), and observed power was maximal (1.0).

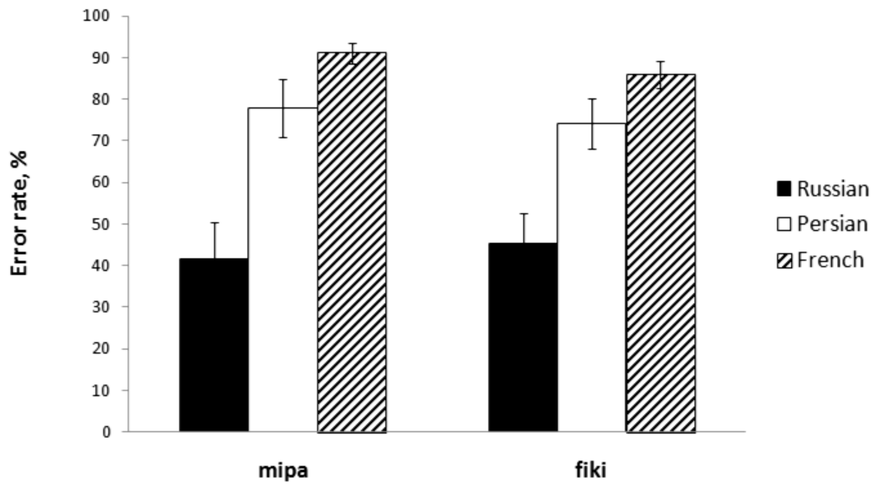

Figure 3. Sequence recall task Mean Error Rate (%) across all language groups

Because of the unequal sample sizes in the three groups, a Dunn-Sidak multiple comparison procedure was performed. The results demonstrated that Russian participants were significantly more accurate in their ability to encode and record stress patterns than either of the two other groups ($p < 0.01$). Although Persian speakers again performed better than French speakers, the difference in error rate did not reach statistical significance. The results of the two tasks are analyzed in the discussion section.

5. Discussion

The purpose of the study was to establish whether there is an effect of L1 metrical system on processing of nonnative word stress contrasts. Specifically, we were interested in determining whether speakers of Russian, Persian and French–languages with different metrical systems–experience unequal processing difficulties related to word stress perception.

The assumption made in this study was that if the learner is not able to perceive L2 stress correctly in the first place, the input might not act as a triggering factor, and the relevant parameters will not be reset. Hence, acquisition of L2 word stress might pose an insurmountable problem for such category of language learners. Conversely, if the perception problem is overcome, the learner should be able to establish appropriate phonological representations for L2 stress and reset necessary parameters. The learner, however, might or might not succeed in doing so, and so may get stuck at the interlanguage stage without ever attaining the correct L2 metrical grammar. Finally, for L1 speakers whose word stress parameters are close to those in L2, perception of word stress should not be a problem, and parameter adjustment might proceed “painlessly”.

The results of the present study partly conform to the original predictions. Thus, it was found that Russian speakers outperform Persian and French speakers on all tasks and process prosodic contrast under investigation significantly faster and more accurately. According to the interpretation of the existing models of stress acquisition, for Russian speakers, stress appears to be a non-detachable aspect of the phonological information, so they are able to encode it at an abstract mental level and use automatic coding as a strategy while making an identification decision. In the present study, Russian participants were able to cope with the stress perception task despite having to rely on nonnative acoustic cues rather than those used in their L1, which indicates that they might have the necessary parameters in place to be able to do that.

Opposite Russian speakers are French participants, who performed quite poorly on both tasks in the present study. This observation is consistent with earlier studies and confirms the presence of stress ‘deafness’ effect in French speakers. In French, stress carries no lexical meaning, but serves a demarcating function. For this reason, French speakers do not have to store stress in their mental representation of a word. Instead, French speakers might process L2 stress by relying on acoustic strategies, such as acoustic mismatch strategy or explicit categorization strategy (Dupoux et al., 2001). However, these strategies inevitably fail when it comes to speeded processing and tasks with higher memory load. An important implication of such persistent stress perception problem concerns ultimate L2 stress acquisition. We can speculate that L2 learners are capable of resetting (or not) a parameter that has a certain value in their L1 to another value in L2 (Archibald, *passim*). However, when a certain representation is absent from L1 (like a representation of word stress in French) and the parameter has never been activated in L1, it is questionable whether L2 is able to generate a new structure, i.e. create a new parameter from scratch as opposed to just changing its value.

The main finding of the present study, however, concerns Persian speakers. In Persian, word stress is quite predictable, but there are also some exceptions to the rule; therefore, Persian stress can also play a contrastive role in order to distinguish different parts of speech. This means that stress information should be represented in the mind of Persian speakers at some abstract level. Based on these assumptions, we hypothesized that Persian speakers should constitute a transitory group between Russian and French speakers. Our predictions were partially borne out. Indeed, the data obtained from the identification task and the speeded sequence recall task display such a trend, with Persian speakers performing better than French speakers, but worse than Russian speakers. However, such difference was only significant in the error rate analysis for the identification task. In light of the reviewed theories, two explanations are possible.

On the one hand, we could propose that stress 'deafness' is not dichotomous and language-specific, and that it is represented in some speakers to a greater extent than in others despite the L1 metrical differences. In this scenario, earlier findings about stress 'deafness' in French speakers can be extended to Persian speakers. But then, the alleged perceptual insensitivity to contrastive stress patterns in Persian speakers cannot be due to the lack of a contrastive property in the speaker's lexicon as had been previously proposed (Kijak, 2009), because Persian stress does play a contrastive role in certain cases.

Instead, we suggest that stress processing difficulties observed in Persian speakers might have a different source. Our Persian participants might still be able to decode and store stress at an abstract phonological level in L1, but the relevant parameters might not be in place when it comes to processing nonnative stress patterns. According to our results, Persian speakers are more accurate in identifying the correct stress than French speakers, but not fast enough (as exemplified by the reaction time data and the speeded sequence recall task data). Thus, for Persian speakers, the problem might not be the perception problem, but the grammar learning problem. Along the same lines, Obler (1988) proposed that parameter setting might be viewed as a gradual process in terms of automaticity. Over the course of repetition, the parameter might get set, and, consequently, speech processing might become faster and more efficient. From this perspective, Persian participants' poor scores might be the result of the lack of automaticity due to the incomplete parameter resetting and the limited exposure to the contrasts in question.

In order to be able to tell where exactly the processing problems originate, it would be interesting to compare the developmental pattern of stress perception by Persian and French speakers provided they are given appropriate amount of perceptual training and see whether further practice and exposure to contrasts under consideration make a difference for these two categories of learners. Resistance to training might be indicative of the perception problem, but improvement in performance might point to grammar learning and gradual parameter resetting.

In conclusion, it is necessary to emphasize that our results are based on a small sample of participants; therefore, the findings of this study should be empirically verified with larger populations in the future. Moreover, the present study employed two tasks, each of which has its own advantages and disadvantages. While sequence recall task reduces a potential of response bias and seems to be good at showing how stress processing depends upon a combination of memory load and phonetic variability (Dupoux et al., 2001), it does not measure reaction times. The identification task, on the other hand, is relatively easy to administer, measures reaction time and error rate, but recency effects due to memory become a consideration. Thus, we suggest that different tasks and both reaction time and error rate measures are used to investigate word stress processing. Finally, in order to obtain a comprehensive understanding of how stress is processed and learned, further research needs to be extended to speakers of other languages.

6. Conclusion

As has been widely documented in the SLA literature, conscious metalinguistic knowledge of an L2 feature or structure does not guarantee a fully successful automatic processing. There are factors that interfere with this process, and L1 influence is one of them. The results of the present study are consistent with previous research findings, and speak to the fact that L1 effects are not restricted to processing of nonnative sounds only, but extend to higher-order units, such as prosody. Therefore, a further investigation and a better understanding of L1 transfer effects with regard to processing and acquisition of L2 prosody is important and might shed new light on the intricate and interesting interactions of phonetic and prosodic information and inform the existing models of speech perception and production.

References

Altmann, Heidi (2006). The perception and production of second language stress: A cross-linguistic experimental study. Doctoral Dissertation, University of Delaware.

Archibald, John (1993). *Language learnability and L2 phonology: The acquisition of metrical parameters*. Dordrecht/Boston/London: Kluwer Academic.

Archibald, John (1995). The acquisition of stress. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp.81-109). New Jersey: L. Erlbaum Associates Inc.

Archibald, John (1997). The acquisition of English stress by speakers of nonaccentual languages: lexical storage versus computation of stress. *Linguistics, 35*, 167-181.

Archibald, John (1998*). Second language phonology*. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Boersma, Paul & Weenink, David (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.08.

Cohen, Jacob (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Cutler, Anne & Mehler, Jacques (1993). The periodicity bias. *Journal of Phonetics, 21*, 103–108.

Dresher, Elan & Kaye, Jonathan (1990). A computational learning model for metrical phonology. *Cognition, 34*(2), 137-195.

Dupoux, Emmanuel, Pallier, Christophe, Sebastián-Gallés, Núria & Mehler, Jacques (1997). A destressing 'deafness' in French? *Journal of Memory and Language, 36*, 406-421.

Dupoux, Emmanuel, Peperkamp, Sharon, & Sebastián-Gallés, Núria (2001). A robust method to study stress 'deafness'. *Journal of the Acoustical Society of America, 110*, 1606-1618.

Dupoux, Emmanuel, Sebastian-Galles, Núria, Navarrete, Eduardo & Peperkamp, Sharon. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*(2), 682-706.

Ferguson, Charles (1957). Word Stress in Persian. *Language, 33*(2), 123-135.

Halle, Morris & Vergnaud, Jean-Roger (1987). Stress and the cycle. *Linguistic Inquiry, 18*, 45-84.

Idsardi, William J. (1992). The computation of prosody. *Unpublished doctoral dissertation*. MIT.

Kijak, Anna (2009). How stressful is L2 stress? A cross-linguistic study of L2 perception and production of metrical systems. Doctoral Dissertation, Universiteit Utrecht.

Liberman, Mark (1975). *The intonational system of English*. Doctoral Dissertation, MIT.

Liberman, Mark & Prince, Alan (1977). On stress and linguistic rhythm. *Linguistics Inquiry, 8*, 249–336.

Mahjani, Behzad (2003). An instrumental study of prosodic features and intonation in modern Persian. MS Dissertation, University of Edinburgh.

Nguyen, Anh-Thu T. & Ingram, John (2005). Vietnamese acquisition of English word stress. *TESOL Quarterly, 39*(2), 309–319.

Nguyen, Anh-Thu T., Ingram, John & Pensalfini, Rob (2008). Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics, 36*, 158–190.

Obler, Loraine (1988). Neurolinguistics and parameter setting. In S. Flynn and W. O'Neil (Eds.), *Linguistic theory in second languages acquisition* (pp. 117-25). Dordrecht: Kluwer Academic Publishers.

Pater, Joe (1997). Metrical parameter missetting in second language acquisition. In S. J. Hannahs & M. Young-Scholten (Eds.), *Focus on phonological acquisition* (pp. 234–261). Amsterdam: John Benjamins.

Peperkamp, Sharon & Dupoux, Emmanuel (2002). A typological study of stress 'deafness'. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology, 7* (pp. 203-240). Berlin : Mouton de Gruyter.

Ringbom, Håkan (1990). Effects of transfer in foreign language learning. In H. Dechert (Ed.), *Current trends in European second language acquisition research multilingual matters* (pp. 205-218). Clevedon.

Roca, Iggy (1999). Stress in the Romance languages. In H. Van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (pp. 659-811). Berlin: Mouton de Gruyter.

Speer, Shari & Ito, Kiwako (2009). Prosody in first language acquisition - Acquiring intonation as a tool to organize information in conversation. *Language and Linguistics Compass, 3*(1), 90-110.

Svetozarova, Natalia (1998). Intonation in Russian. In D. Hirst, & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp 261-274). England: Cambridge U Press.

Swan, Michael & Smith, Bernard (2001). *Learner English: a teacher's guide to interference and other problems* (2nd ed.). Cambridge University Press.

Tremblay, Annie & Owens, Nathan (2010). The role of acoustic cues in the development of (non-)target-like L2 prosodic representations. *Canadian Journal of Linguistics, 55*, 85-114.

Van der Pas, Brigit & Zonneveld, Wim (2004). L2 Parameter resetting for metrical systems. *The Linguistic Review, 21*(1), 125-170.

Windfuhr, Gernott (1990). Persian, In Comrie Bernard (Ed.), *The world's major languages*. New York: Oxford.

Selected Proceedings of the 2010 Second Language Research Forum: Reconsidering SLA Research, Dimensions, and Directions

edited by Gisela Granena, Joel Koeth, Sunyoung Lee-Ellis, Anna Lukyanchenko, Goretti Prieto Botana, and Elizabeth Rhoades

Cascadilla Proceedings Project Somerville, MA 2011

Copyright information

ISBN 978-1-57473-448-5 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Lukyanchenko, Anna, William J. Idsardi, and Nan Jiang. 2011. Opening Your Ears: The Role of L1 in Processing of Nonnative Prosodic Contrasts. In *Selected Proceedings of the 2010 Second Language Research Forum*, ed. Gisela Granena et al., 50-62. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2615.