LANGUAGE LEARNING
*A Journal of Research in Language Studies*

**Language Learning ISSN 0023-8333**

# Context Effects in the Processing of Phonolexical Ambiguity in L2

Anna Chrabaszcz and Kira Gor

University of Maryland

In order to comprehend speech, listeners have to combine low-level phonetic information about the incoming auditory signal with higher-order contextual information to make a lexical selection. This requires stable phonological categories and unambiguous representations of words in the mental lexicon. Unlike native speakers, second language (L2) speakers, who perceive nonnative sounds through the prism of their first language (L1), operate with fuzzy phonological categories, which lead to phonologically ambiguous lexical representations (e.g., the words *rock* and *lock* can be confused if phonological representations for /r/ and /l/ are not sufficiently robust). The present study uses the AX discrimination task to establish the degree of sensitivity of L2 listeners to the Russian hard/soft phonological contrast. The same phonological contrasts are then used in the stimuli for the second task—listening comprehension task with word identification—to mark semantic, syntactic, and morphological distinctions in words. The goal of the study is to examine the contributions and relative efficiency of different contextual constraints (semantic, syntactic, and morphological) to the resolution of phonolexical ambiguity in L2 auditory sentence processing. The results suggest that when L2 phonological contrasts present a discriminability problem and create phonolexical ambiguity, L2 listeners rely on morphological constraints for disambiguation of word forms and syntactic constraints for disambiguation of words belonging to different parts of speech to a greater extent than on semantic constraints for disambiguation of nouns in the same form.

**Keywords** phonolexical ambiguity; word identification; context bias; L2; Russian; hard and soft consonants

We would like to thank the anonymous reviewers who evaluated this manuscript for their thoughtful comments and suggestions. Any remaining errors are our own. The study was supported in part by the School of Languages, Literatures and Cultures, and the College of Arts and Humanities at the University of Maryland.

Correspondence concerning this article should be sent to Anna Chrabaszcz (Lukianchenko), 3215 Jimenez, School of Languages, Literatures, and Cultures, University of Maryland, College Park, MD 20742. E-mail: lav@umd.edu



DOI: 10.1111/lang.12063

## Introduction

When coping with, for example, noise in speech and having to resolve perceptual ambiguities in listening, both first language (L1) and second language (L2) users struggle with extracting information at the phonetic-phonological level and must use some compensation mechanisms to disambiguate word meaning. While L1 comprehenders can rely on context to disambiguate meaning of spoken words, it is unclear whether L2 comprehenders have access to the same mechanisms, owning to their (i) smaller vocabularies and weaker semantic associations between words, (ii) use of shallow syntactic processing and decreased sensitivity to morphosyntactic violations, and (iii) slowed meaning integration and prediction processes.

Although numerous studies, to be reviewed later, have used different methodologies to identify the perceptual difficulties that L2 listeners face in speech perception as well as to examine how these perceptual deficits affect lexical processes, little is known about how these difficulties play out at sentence level of speech comprehension, and how information derived from the perceptual analysis interacts with information derived at higher levels of linguistic analysis. The present study takes on the exploratory research goal of examining how different contextual cues (semantic, morphological, and syntactic) contribute to the identification of L2 phonologically ambiguous words. By tackling these issues, the study pursues the broader goal of shedding light on what distinguishes native versus nonnative speech comprehension.

## L2 Speech Perception

Unlike native speech perception, which is robust, automatic, and efficient even in nonoptimal conditions, L2 speech perception is notoriously problematic even in highly proficient and experienced L2 listeners. Phonetic segments that are phonologically contrastive in L2, but not in L1, are often miscategorized and misconstrued by L2 speakers, which renders spoken L2 comprehension difficult (Strange & Shafer, 2008). Perhaps the best-known documented evidence in the second language acquisition (SLA) literature is the difficulty that Japanese listeners experience with the discrimination of the English /r/ and /l/ contrast, as in *rock* versus *lock*, because they are perceived as allophonic variations of the same phoneme in Japanese (Goto, 1971; McClelland, Thomas, McCandliss, & Fiez, 1999; Miyawaki et al., 1981). Another well-attested example concerns the perceptual difficulty of the Catalan /e/-/ɛ/ contrast for

Spanish-dominant Spanish-Catalan bilinguals. Unlike Spanish, Catalan has two mid vowels with different heights, one high /e/ and one low /ɛ/, which are used to distinguish between words, for example, /te/ 'take' and /tɛ/ 'tea.' Using a variety of research methods and experimental paradigms, Spanish-dominant bilinguals were systematically demonstrated to perform rather poorly on the tasks involving this Catalan contrast compared to the control group of Catalan-dominant speakers (Bosch, Costa, & Sebastián-Gallés, 2000; Navarra, Sebastián-Gallés, & Soto-Faraco, 2005; Pallier, Bosch, & Sebastián-Gallés, 1997; Pallier, Colomé, & Sebastián-Gallés, 2001; Sebastián-Gallés, Rodríguez-Fornells, de Diego-Balaguer, & Díaz, 2006; Sebastián-Gallés & Soto-Faraco, 1999). This evidence is particularly striking because such perceptual discrimination difficulty is observed even in highly proficient bilinguals, who receive an early and extensive exposure to Catalan and who use both languages in their everyday life.

What are the consequences of this perceptual problem? According to Broersma and Cutler (2011), listeners suffer from the difficulties of sound perception only to the extent that their word recognition and speech comprehension is affected. To illustrate, if L2 listeners experience perceptual difficulties with a certain L2 phonological contrast, the minimal pairs distinguished on the basis of such contrast will become homophonous. For example, the words *rock* and *lock* may sound so similar to Japanese listeners that they can even identify them as the same word. In contrast, when native speakers of English hear the word *rock*, they are able to extract the correct meaning of the word through mapping each sound segment in the word onto a corresponding phoneme and through matching up the resulting phoneme string with the correct lexical candidate. Provided that spoken word recognition is achieved through simultaneous activation of multiple lexical candidates (e.g., Goldinger, Luce, Pisoni, & Marcario, 1992; Marslen-Wilson & Welsh, 1978; McClelland & Elman, 1986; Zwitserlood, 1989) and the subsequent choice of the best-fitting candidate among the competing candidates, robust phonological information can effectively contribute to the lexical search and narrow down the list of possible lexical candidates at early stages of processing (i.e., *rock* will be selected over *lock*, *mock*, *rack*, *rob*, and so on), while imprecise perception can result in lexical ambiguity and lead to increased activation of phonological neighbors and spurious competition (Weber & Cutler, 2004).

To examine the consequences of phonological ambiguity on lexical representations, Pallier et al. (2001) used an auditory repetition-priming paradigm with Spanish-dominant and Catalan-dominant bilinguals. It was found that the

Spanish-dominant listeners, but not the Catalan-dominant listeners, showed repetition-priming effect for minimal pairs involving a phonological contrast distinctive in Catalan but not in Spanish (e.g., /netə/ 'granddaughter' vs. /nɛtə/ 'clean'). The authors concluded that the Spanish-dominant participants perceived these minimal pairs as homophones, thus suggesting that phonological ambiguity resulted in lexical ambiguity. Using a lexical decision task, Sebastián-Gallés, Echeverría, and Bosch (2005) demonstrated that the Spanish-dominant Spanish-Catalan bilinguals tended to accept nonwords created by substituting the Catalan /e/-/ɛ/ contrast as words significantly more often than the Catalan-dominant bilinguals and significantly more often than nonwords with the substitution of the control vowel contrast, /i/-/u/, which is common in both languages. In a subsequent study, Sebastián-Gallés et al. (2006) corroborated their behavioral findings with evidence from an event-related brain potential (ERP) experiment. They found that Catalan-dominant bilinguals and Spanish-dominant bilinguals differed in terms of the elicited error-related negativity (ERN) component. In their study, a clear ERN-like component was expected if participants were fairly sure that their response was erroneous. In contrast, for insecure or doubtful decisions, the differences between correct and error responses were expected to be less evident. The researchers observed that Catalan-dominant bilinguals showed ERN differences between their erroneous responses to manipulated nonwords and correct responses to words in the /e/-/ɛ/ condition, whereas Spanish-dominant bilinguals showed no differences between the two types of responses. In fact, their correct responses to real words and their incorrect responses to the critical nonwords (i.e., when they falsely accepted them as words) showed the same degree of uncertainty. Combined with the behavioral data, the results suggest that Spanish-dominant bilinguals simply tend to accept manipulated nonwords as real Catalan words most of the time.

Similarly to that in bilinguals, phonological difficulties were also demonstrated to have consequences for lexical representations in late L2 learners. A series of experiments examined recognition of L2 English spoken words by L1 Dutch listeners, which was assumed to be contingent on the listeners' ability to discriminate between confusable phonemes in the L2 (Broersma, 2005; Broersma & Cutler, 2011; Díaz, Mitterer, Broersma, & Sebastián-Gallés, 2012). Dutch lacks the English /æ/-/ɛ/ vowel distinction, which could lead to the word *flash* being interpreted as *flesh* by Dutch speakers of English. This is exactly what the findings demonstrated. In auditory lexical decision tasks, Dutch listeners accepted nonwords (e.g., *lemp*) as real English words (e.g., *lamp*) more often than English listeners did. In a cross-modal priming task, nonwords extracted from word or phrase contexts (e.g., *lemp* from *eviL EMPire*) led to increased activation of corresponding real words (*lamp*) for Dutch, but not for native speakers of English.

To summarize, numerous studies have used different methodologies to identify the perceptual difficulties that L2 listeners face in speech perception as well as to examine how these perceptual deficits affect lexical processes. But little is known about how these difficulties play out at sentence level of speech comprehension and how information derived from the perceptual analysis interacts with information derived at higher levels of linguistic analysis.

## Context Effects in L2 Speech Comprehension

When words are presented in isolation, the quality of the acoustic-phonetic input plays a crucial role in word identification. However, words rarely occur in isolation in natural speech. Instead, words are strung together, and the way they are connected in sentences is mediated by complex semantic, syntactic, and morphosyntactic relationships among them. That is why spoken word recognition in continuous speech does not only entail attending to the phonological form, but also engages higher-order processes (e.g., lexical processes, syntactic processes, compositional processes, and so on).

Native speakers have both robust phonological decoding/encoding strategies and effective syntactic and morphosyntactic processing strategies. They also have more experience with the structural properties and distributional patterns of phonemes, words, phrases, and sentences, as well as with the sociocultural schemata evoked by the context in which the communication act is happening. This allows them to take advantage of the higher-order, contextual information in a rapid, effortless, and efficient manner, and build linguistic predictions to speed up computation of incoming words (Craig, Kim, Rhyner, & Chirillo, 1993; Marslen-Wilson & Tyler, 1980; Salasoo & Pisoni, 1985; Tyler & Wessels, 1983). Thus, when native listeners are faced with lexically ambiguous (e.g., the word *bank* referring to either the edge of a river or a financial institution) or perceptually ambiguous input (e.g., due to the background noise or several people speaking at the same time), the lexical processor is able to compensate for the ambiguity by resorting to other sources of information, such as prosodic, pragmatic, semantic, and so on (Gaskell & Marslen-Wilson, 2001; Lucas, 1999; Mirman, 2008; Moss & Marslen-Wilson, 1993; Tabossi & Zardon, 1993).

Unlike native speakers, L2 speakers struggle with extracting information at the phonetic-phonological level reliably and unambiguously. However, it has been proposed that, because L1 and L2 speakers face similar problems in spoken and written word comprehension (e.g., having to cope with noise in speech and having to resolve lexical ambiguities in reading and listening), the mechanisms for using context to compensate for ambiguities should be available to L2 listeners as well (Broersma & Cutler, 2008). Using reaction time measures in conjunction with ERPs, Elston-Güttler and Friederici (2005, 2007) observed that L1 and L2 processing of lexical ambiguities (homonyms) in reading is fundamentally similar: Despite the fact that L2 speakers are slightly slower than natives at context integration processes around 500 ms, during the final processing phase (around 800 milliseconds) disambiguation is comparable in the L1 and the L2. Because phonologically ambiguous words essentially become homonyms for L2 listeners, they should be able to use contextual cues for meaning disambiguation in situations of ambiguous phonetic-phonological information as well.

Let us consider the above example of phonological ambiguity between the words *rock* and *lock* for a hypothetical Japanese learner of English. Such ambiguity may be resolved at a sentence level with the help of lexical-semantic context, in which the word *lock*, for example, has a very low cloze probability (1b) in comparison with the word *rock* (1a).

(1) a. I climbed a **rock** for the first time in my life.
b. I climbed a ***lock** for the first time in my life.

Similarly, syntactic and morphological information can help disambiguate meaning in phonologically ambiguous contexts, as in (2), where a verb is expected to occur after the auxiliary *didn't* (2a), but not an adverb (2b).

(2) a. He didn't **arrive** until noon the next day.
b. He didn't ***alive** until noon the next day.

Although it seems apparent that L2 listeners should make use of contextual cues in order to arrive at the correct meaning interpretation, research on L2 speakers' use of context is very scant and insufficient (except for a handful of studies, e.g., FitzPatrick & Indefrey, 2010; Li & Yip, 1998; Love, Maas, & Swinney, 2003; Sanders & Neville, 2003) to make strong predictions about what kinds of contextual information they benefit from the most. There is some evidence showing that L2 comprehenders rely on broad semantic-lexical context to aid them in compensating for gaps in understanding. Field (2004) cites a classroom-based study, which showed that whenever the learners did not understand certain words in a text, they tended to replace them in their oral productions with similar-sounding words that fit the overall semantic and thematic schema of the text. For example, in a text about travel, one student substituted the word *mat* with *map* and another one used the word *bridge* instead of *ledge*. Motivated by similar observations, Field conducted an experiment where he intentionally substituted highly predictable words at the end of semantically constraining sentences with similar sounding words (e.g., *We arrived at the airport on time, then we had to wait two hours for the* ***train***, where *train* replaced the better-fitting *plane*). The substitute word was always less predictable but nonetheless acceptable in the context. The learners were asked to write down the last word in each sentence. He found that the recovery of the original, better-fitting word ranged from 15% to 62% of the responses, which suggests that L2 speakers are capable of constructing lexical-semantic predictions and use this information for meaning resolution.

A study by Hu and Jiang (2011) used a cross-modal priming task to examine integration of semantic information in sentences by Chinese L2 learners of English and similarly found that L2 listeners benefitted from semantically constrained contexts for word recognition. They were asked to perform a lexical decision task on the targets which were preceded by different sentential contexts: semantically congruent, neutral, and incongruent (e.g., *The girl mailed the letter without a stamp/ sticker/ stone*). L2 listeners showed a facilitation effect only for the semantically congruent condition. Chambers and Cooke (2009) demonstrated that L2 speakers can also use sentential semantic constraints to reduce interlingual lexical activation. They observed that when L2 French listeners heard the target *poule* (‘chicken’) in the nonrestrictive semantic context, their eye movements reflected temporary consideration of a similar-sounding English competitor *pool*. When the context was restrictive and incompatible with the competitor, fixations on the competitor were dramatically reduced.

More evidence on L2 listeners’ use of lexical-semantic context comes from studies on listening to speech in noise, in which researchers consistently find that words in semantically predictable sentence contexts are identified more quickly and accurately than words spoken in isolation or embedded in unpredictable sentence contexts at all noise levels, although the semantic context advantage is greater for L1 than for L2 listeners (Bradlow & Bent, 2002; Cutler, Garcia Lecumberri, & Cooke, 2008). Electrophysiological evidence also suggests that L2 speakers use sentential context to construct lexical-semantic expectations. Thus, they exhibit the same N400 effect for semantic violations in sentences as native speakers, although this effect is sometimes delayed and has a slightly reduced amplitude (Hahne, 2001; Hahne & Friederici, 2001).

With regard to the role of syntactic and morphological cues in meaning disambiguation, empirical studies on this topic are practically nonexistent. Of course, L2 speakers can only rely on morphosyntactic cues provided they have mastered the relevant grammar, but considerable SLA evidence suggests that differences in language experience (such as age of acquisition, length of residency in the L2 country, and type of instruction) have a much more pronounced effect on morphosyntactic rather than semantic L2 processing (e.g., Johnson & Newport, 1989; Weber-Fox & Neville, 1996; see also Hahne & Friederici, 2001). Clahsen and Felser (2006a, 2006b) argue that even though the basic architecture of the processing system is the same in the L1 and L2, shallow morphosyntactic parsing predominates in L2 processing. According to their shallow structure hypothesis, the representations which L2 learners compute during sentence processing contain less syntactic detail than those of native speakers and lack complex hierarchical structure and abstract, configurationally determined elements, which makes learners rely more on nonstructural information in parsing ambiguous sentences (Felser, Roberts, Marinis, & Gross, 2003; Papadopoulou & Clahsen, 2003).

Furthermore, electrophysiological components associated with morphosyntactic processing have been shown to be weakened or absent in L2 speakers. In studies by Hahne (2001) and Hahne and Friederici (2001), L1 Russian and L1 Japanese learners of L2 German exhibited an N400 response to semantic violations similar to that of the native speakers, but neither L2 group showed an early left anterior negativity (ELAN) response (see, however, Rossi, Gugler, Friederici, & Hahne, 2006, for contradictory findings). The P600 was reduced for the Russian speakers and was absent in the Japanese speakers of German. It has been suggested that the lack of early negativity components, such as ELAN, in L2 speakers indicates that they do not employ the same early, highly automatic syntactic parsing mechanisms as native speakers. The absence of the P600 effect, on the other hand, is indicative of the L2 speakers' difficulty with late syntactic repair processes. It should be noted, however, that electrophysiological responses in L2 populations have been shown to be strongly modulated by L2 speakers' proficiency (e.g., Rossi et al., 2006).

Taking the above evidence into account, it is not clear to what extent L2 listeners will be able to benefit from semantic and morphosyntactic contextual constraints to disambiguate the identity of words during speech comprehension. If they are more sensitive to the lexical-semantic content of the utterance (as the literature suggests), they should be relying on semantic cues to resolve phonologically ambiguous words. If, however, they pay more attention to morphosyntactic cues, those should prevail in word disambiguation. Because no previous study has investigated the effect of different types of contextual constraints on phonological ambiguity resolution within the same experimental setup, the present study takes on the exploratory research goal of examining how different contextual cues (semantic, morphological, and syntactic) contribute to the identification of L2 phonologically ambiguous words.

## The Present Study

The present study is based on a set of beliefs about what distinguishes native versus nonnative speech comprehension. These beliefs are centered around the idea that phonology plays a crucial role in lexical processes (such as lexical activation, selection, and integration) and that fuzzy, unclear phonological perception in the L2 renders spoken comprehension problematic by making similar-sounding words highly confusable (or even homophonous), thereby creating lexical ambiguity. We refer to such ambiguity as phonolexical ambiguity. It is not simply phonological ambiguity, because it affects lexical representations. And it is not precisely lexical ambiguity (e.g., *bank* referring to the bank of the river and a financial institution) the way it is understood in the literature on L1 ambiguity resolution. In the case of phonolexical ambiguity, L2 listeners do not deal with pure homonyms (like *bank*) with two possible meanings; they are also probably aware of the orthographic differences underlying the confusable word pair (e.g., *rock* and *lock*) and can distinguish the confusable words in writing. They may even hear some acoustic differences between the two words when they are pronounced in isolation, but imperfect phonological encoding strategies may make it difficult to decide between the two alternatives and may lead to spurious lexical competition when it comes to the selection of the needed lexical candidate during speech processing.

Although phonolexical ambiguity in L2 has been explored in earlier studies (e.g., Pallier et al., 2001; Sebastián-Gallés et al., 2005), these have all focused on establishing its consequences for the lexical processor at the word level. To the best of our knowledge, no SLA study has examined how such ambiguity plays out at the sentence level and how L2 listeners deal with it.

In sentence comprehension, whether in the L1 or L2, low-level information about phonemes and phonotactics needs to be combined with high-level (e.g., semantic, syntactic, and so on) and wider schematic information via bottom-up and top-down processing mechanisms in order to make speech comprehension possible. The review of the literature presented above suggests, however, that L2 listeners may not always possess the necessary lexical-semantic knowledge (e.g., their lexicon is smaller, the lexical-semantic associations are weaker) and/or morphosyntactic knowledge which they would be able to utilize spontaneously and effortlessly for comprehension of aural texts. That is why the question of whether L2 speakers can rely on contextual information to help them resolve phonolexical ambiguities in a way that is akin to L1 speakers becomes a nontrivial one.

The main contribution of the present study, despite its exploratory nature, is to fill the gap in the existing literature by quantifying context effects (semantic, morphological, and syntactic) in the processing of phonolexical ambiguity at the sentence level with the goal of finding out which cues are the strongest in L2 listening. In addition, this study offers insights on the role of phonology and perceptual difficulty in L2 listening comprehension and the way it interacts with syntactic, morphological, and semantic information.

First, we conducted a phonetic discrimination task to evaluate L2 listeners' perceptual difficulty with the target phonological contrasts in the absence of any contextual information. Next, we carried out a listening comprehension task, in which listeners had to discriminate words in sentences on the basis of the same phonological contrasts. The words were embedded in either contextually congruent or incongruent sentences. Such a design enabled us to estimate whether L2 listeners' word choice is predominantly determined by low-level, phonetic information or whether the information is coming from contextual sources. The hypothesis was that, if L2 listeners experience perceptual difficulty discriminating target contrasts on the basis of low-level bottom-up information in the discrimination task, this difficulty should carry over to the listening task and should affect participants' discrimination of the minimal pairs of real words with the same contrasts. Therefore, if L2 listeners primarily rely on bottom-up information in sentence processing, their ability to discriminate the target contrast should directly predict their word identification performance on the listening task, independent of the type of contextual constraint (semantic, morphological, or syntactic). In other words, their word discrimination performance should be the same in all context conditions. If, however, higher-order contextual cues are available to L2 listeners during sentence processing and some contextual cues are more effective than others in helping to disambiguate ambiguous words in sentences, L2 listeners should demonstrate unequal context bias in those conditions.

## Method

### Participants

Forty-four participants took part in the study. The first group served as a control group and included twelve native speakers of Russian (1 male, 11 female). Their age ranged from 22 to 60 years old ($M = 33.9$). Most of them were students in a mid-Atlantic university in the United States; others were recent university graduates and were working in the United States at the time of testing. One person who was 60 years old was a visiting professor from Russia teaching in a U.S. university at the graduate level. All of the native Russian participants had spent most of their life in Russia.

The L2 group included 32 American learners of Russian (26 male, 6 female). Most of the participants were enrolled in intensive Russian programs, and some had spent 2 years working in Russia or Russian-speaking countries. Several participants had Russian spouses. One participant indicated that he did not have formal instruction in Russian, but had learned the language primarily via immersion while living in Russia. Nine participants had taken official proficiency tests in Russian before, such as the Oral Proficiency Interview and the Defense Language Proficiency Test, and scored above intermediate high.

All L2 speakers were asked to complete a language background questionnaire with questions about their schooling, language learning experience and language use as well as to rate their Russian proficiency on a scale from 1 (minimum) to 10 (maximum) in different linguistic domains. In addition, prior to the experiment, all L2 participants were required to take a Russian cloze test developed and piloted at the University of Maryland. The test takers had to read the text with 25 blanks and supply only one word per blank. They were given 15 minutes to complete the test. The maximum number of correct responses was 25. Each accurate—exact or acceptable—response received a score of 1; otherwise a score of 0 was given. The scores were added up to yield a final overall proficiency score. The language background information as well as participants' proficiency scores are presented in Table 1.

### Target Phonological Feature

The phonological feature used in this study was consonantal hardness/softness (also sometimes referred to as velarization/palatalization) in Russian. The distinction between hard and soft consonants in Russian is very difficult to master for English-speaking learners of Russian (Bondarko, 2005; Diehm, 1998; Lukyanchenko & Gor, 2011). It permeates almost the entire consonantal system in Russian and is used contrastively, for example, ***вес*** ('weight', /v$^j$es/)—***весь***

**Table 1** L2 participants' language background information

| | Mean | *SD* | Range | |
|---|---|---|---|---|
| | | | Min | Max |
| Age | 24 | 4.02 | 20 | 41 |
| Age when started learning Russian | 18.75 | 1.97 | 14 | 24 |
| Age when traveled to a Russian-speaking country | 19.59 | 1.76 | 17 | 25 |
| Length of living in a Russian-speaking country (years) | 1.71 | 0.52 | 0.13 | 2 |
| Formal instruction in Russian (years) | 2.39 | 1.81 | 0 | 7 |
| Percent of daily use of Russian at home (%) | 6.72 | 10.34 | 0 | 50 |
| Percent of daily use of Russian at work (%) | 20.34 | 20.69 | 0 | 80 |
| Self-rated pronunciation in Russian | 6.97 | 1.40 | 4 | 9 |
| Self-rated listening proficiency in Russian | 7.94 | 1.76 | 1 | 10 |
| Self-rated reading proficiency in Russian | 7.22 | 1.90 | 1 | 10 |
| Self-rated writing proficiency in Russian | 6.44 | 1.54 | 3 | 9 |
| Self-rated knowledge of Russian grammar | 7.03 | 1.62 | 2 | 10 |
| Cloze test (Proficiency measure) | 18.91 | 4.83 | 14 | 25 |

('the whole of', /v$^j$es$^j$/). Usually, soft consonants are interpreted as the ones having a secondary articulation (the raising of the middle part of the tongue toward the hard palate) compared to the corresponding hard ones. In consonant-vowel (CV) syllables, the feature flows at the syllable level as well as at the segmental level, that is, the process of CV accommodation smears featural information and distributes it over the syllable. Importantly, while in English the effect of the vowels on the consonants is usually greater, the opposite seems to be true for Russian: The consonants are more stable and independent of the vowels, and it is the vowels that accommodate themselves to the consonants through coarticulation (Howie, 2001). Therefore, most salient cues to the softness of a consonant in a CV syllable in Russian are mainly contained in the formant transitions of the following vowel (Bondarko, 2005; Kochetov, 2002). If the consonant is not followed by a vowel (e.g., when it occurs in the word-final position), it presents a greater discriminability difficulty for nonnative speakers of Russian (Lukyanchenko & Gor, 2011).

## Task 1: AX Discrimination

### *Procedure*

The purpose of the AX discrimination task was to measure the level of perceptual difficulty of the hard/soft consonant contrasts that differentiated the target words in the listening comprehension task with word identification, namely, coronal obstruents /t-tʲ/, liquids /l-lʲ/, and labial obstruents /f-fʲ/. Participants were presented with pairs of Russian nonce syllables and were instructed to press the keyboard button labeled "same," when the two syllables sounded the same (regardless of the speakers' voice differences), or "different," when the two syllables in the pair sounded different. The interval between the two stimuli in a pair was kept at 100 milliseconds. The response was timed from the end of the second syllable in the pair. Timeout was set at 7 seconds. If participants did not respond within 7 seconds, the next item was presented and the response was considered timed out. Feedback on the accuracy of the participant's responses was given only during the practice stage. No feedback was provided during the test. The AX task was given to participants after the listening task so that the target contrasts were not so obvious and conspicuous in the listening task.

*Materials*

The hard/soft consonant contrast was tested in word-final (coda) position. Six nonce minimal pairs of the following types were constructed: (1) with the hard/soft consonant preceded by the low back vowel /a/: /at-atʲ, al-alʲ, af-afʲ/ and (2) with the hard/soft consonant preceded by the high front vowel /i/: /it-itʲ, il-ilʲ, if-ifʲ/. Each of the six consonant pairs was repeated eight times (four times as an AX sequence, and four times as an XA sequence, where A is a syllable with a hard consonant, and X is a syllable with a soft consonant). The same number of AA- and XX-type sound pairs was also constructed. In total, the stimulus pairs required 48 "same" and 48 "different" responses. In addition, control items were designed (32 same and 32 different pairs). These included nonce minimal pairs with relatively easy contrasts, such as /tʃ-k/, /s-ʃ/, which differed in the manner and/or place of articulation. The critical items were randomized and mixed with control items. Five practice trials were given at the beginning of the task to familiarize participants with the test procedure. All stimuli were recorded by two female talkers, both native speakers of Russian. The same voice was never repeated within the same trial. The stimuli were recorded several times by each talker, and the most unambiguous tokens were selected for the task.

### Task 2: Listening Comprehension With Word Identification

*Procedure*

The listening comprehension task with subsequent word identification allowed us to estimate the extent to which participants rely on information derived at the phonological level and information derived at higher levels of processing, such as semantic, morphological, and syntactic. Participants were randomly assigned to one of the four presentation lists (A, B, C, or D) so that no participant was exposed to the same sentence more than once. On each trial, participants were presented with an auditory sentence for comprehension followed by two words presented on the computer screen. Half of the sentences were congruent and half of the sentences were incongruent. In congruent sentences, the target word fit the context, that is, phonological information aligned with contextual information (e.g., *My younger brother and elder* ***sister*** *are coming to see me tomorrow*). Incongruent sentences contained a mismatch that marked a semantic ('sister/system'), morphological ('seen/sees'), or syntactic category ('seam/seize') violation (e.g., *My younger brother and elder* ****system*** *are coming to see me tomorrow*). At the word identification stage, participants had to decide which of the two words (i.e., *sister* or *system*) occurred in the sentence they had just heard. They pressed LEFT SHIFT button if they chose the word appearing on the left, and RIGHT SHIFT button if they chose the word appearing on the right. The appearance of the words on the computer screen (right or left) was counterbalanced across lists. After word identification decision was made, participants received feedback on the accuracy of their response and were asked to respond to a Yes/No question (presented visually on the screen) about the content of the sentence. Follow-up questions (21 Yes- and 21 No-questions) were included in order to ensure that participants were attending to meaning and not just randomly pressing buttons. At the beginning of the experiment, five practice items were included to familiarize the participants with the test routine and button pressing. After completing the listening task, L2 participants were asked to translate all the words in the Critical condition from Russian into English in order to test their knowledge of the stimulus items.

*Materials*

Four presentation lists with 90 sentences each (360 sentences total) and 42 comprehension questions per list were constructed. The sentences were divided into three conditions: Critical, Control, and Filler. A female native speaker of Russian recorded all sentences. For a complete list of minimal pairs used in this task, see the Appendix.

The sentences in the Critical condition had the target words embedded in the middle of the sentence ($k = 30$). The targets were the minimal pairs differentiated on the basis of the phonological distinction of consonantal hardness/softness in the word-final position, for example, ***мат*** ('stalemate', /mɑt/) versus ***мать*** ('mother', /mɑt$^j$/). Three phonological contrasts were used in the target words in the Critical condition: /t-t$^j$/ (19 word pairs), /l-l$^j$/ (10 word pairs), and /f-fʲ/ (1 word pair). Due to the natural constraints of finding minimal pairs with desired specifications (e.g., having a hard/soft consonant distinction word-finally and avoiding very low-frequency words), the number of word pairs per phonological contrast was unequal. The minimal pairs did not always share the same orthographic representation, but they were always phonologically identical except for the final consonant, for example, ***балет*** ('ballet', /bɑlʲet/) versus ***болеть*** ('to be sick', /bɑlʲetʲ/).

The Control condition consisted of sentences with the minimal pairs that could be easily discriminated on the basis of the consonant's place or manner of articulation ($k = 30$). For example, in ***суд*** ('court', /sut/) and ***суп*** ('soup', /sup/), the word-final consonant is easily differentiated by the place of articulation in both English and Russian. Similarly to the Critical condition, all minimal pairs in the Control condition differed in the word-final phoneme. The difference between the two conditions was in the level of hypothesized perceptual difficulty of the phonological contrast in question. Because it has been shown that [± soft] contrasts occurring in the word-final position present a high level of difficulty for learners of Russian (Bondarko, 2005; Kochetov, 2002; Lukyanchenko & Gor, 2011), we expected word identification in the Critical condition to be more difficult than in the Control condition.

Thirty Filler sentences ($k = 30$) were also added to the test in order to prevent the participants from developing response strategies, for example, listening for the word-final phoneme. Unlike the words in the Critical and Control conditions, where the phonological contrast was position controlled, minimal pairs in the Filler condition differed from each other in one phoneme in any word position except at the end of the word. The contrasting phonemes were either vowels (***стол*** ('table', /stol/) – ***стул*** ('chair', /stul/)) or consonants (***работа*** ('work', /rɑbotɑ/) – ***забота*** ('care', /zɑbotɑ/)). All target words in the Critical and Control conditions were embedded in the middle of 8–12 word phrases, but never occurred at the beginning or at the end of the sentence to avoid the confounding short-term memory effect. The target words in the Filler condition had, however, a more flexible position in the sentence and could occur at the edges of the sentences in order to prevent the listeners from attending only to the middle segment of the sentence. The overall design of the experiment is presented in Table 2.

Furthermore, all sentences in the Critical and Control conditions were subdivided into three subconditions: Semantic, Syntactic, and Morphological. In the Semantic condition, both members of the minimal pairs belonged to the same syntactic category (noun) and shared the same morphological form

**Table 2** Design of the listening comprehension task with word identification

| Condition | Word position | Perceptual difficulty | Contextual constraint: Semantic | Syntactic | Morphological |
|---|---|---|---|---|---|
| Critical, [± soft] | word-final | difficult | 10 | 10 | 10 |
| Control | word-final | easy | 10 | 10 | 10 |
| Filler | any word position | easy | 30 | n/a | n/a |

(singular number, accusative, or nominative case), but differed in terms of semantic congruency: While one word was semantically congruent, its minimal pair counterpart was semantically incongruent in a given sentence, and vice versa (see Table 3 for an example). In the Syntactic condition, members of the minimal pairs belonged to different syntactic categories (e.g., a noun and a verb), as in ***брат*** ('brother', /brɑt/) versus ***брать*** ('to take', /brɑtʲ/). The syntactic structure of the sentence in this condition unambiguously specified the context-appropriate syntactic category assignment for the target word; therefore, substituting one word for the other made the sentence not only semantically, but also syntactically incongruent. Following the same pattern, the Morphological condition included the sentences in which morphosyntactic cues indicated what morphological form of the word should be used (e.g., a verb infinitive after a modal verb). The phonological distinction between the minimal pairs in this condition marked the morphological distinction between a verb in the third person singular and a verb infinitive in the Critical condition, as in ***говорит*** ('speaks', /gɑvɑrʲit) versus ***говорить*** ('to speak', /gɑvɑrʲitʲ/), and a verb in the past tense and a verb infinitive in the Control condition, as in ***сказал*** ('said', /skɑzɑl/) versus ***сказать*** ('to say', /skɑzɑtʲ/). In the Critical condition, the Semantic condition included noun minimal pairs with any of the three tested contrasts (coronal obstruents /t-tʲ/, liquids /l-lʲ/, and labial obstruents /f-fʲ/); the Syntactic condition included minimal pairs with the /t-tʲ/ and /l-lʲ/ contrasts, and the Morphological condition consisted of minimal pairs with the /t-tʲ/ contrast only.

The items were constructed in such a way that congruent target words were highly predictable from the context, although they did not exclude alternative possibilities (e.g., *She always visited her father and* ***mother*** *(or brother, sister,*

**Table 3** An example of a test trial (Critical Semantic condition) across four presentation lists

| Congruency | List | Preceding context | Target word | Following context |
|---|---|---|---|---|
| C | A | Маленький мальчик нарисовал прямой | **угол**/ugɑl/ | в тетради по геометрии |
| | | *A little boy drew a straight* | *angle* | *in his geometry notebook* |
| I | B | Маленький мальчик нарисовал прямой | ∗**уголь**/ugɑlʲ/ | в тетради по геометрии |
| | | *A little boy drew a straight* | ∗*coal* | *in his geometry notebook* |
| C | C | Россия продает газ и каменный | **уголь**/ugɑlʲ/ | странам Европы и Азии |
| | | *Russia sells gas and mineral* | *coal* | *to Europe and Asia.* |
| I | D | Россия продает газ и каменный | ∗**угол**/ugɑl/ | странам Европы и Азии |
| | | *Russia sells gas and mineral* | ∗*angle* | *to Europe and Asia* |

*Note.* C stands for Congruent condition; I stands for Incongruent condition.

*grandparents, etc.) on weekends and holidays)*. To ensure that the target words had on average comparable cloze probabilities across different conditions, a cloze test was conducted, in which 10 native speakers of Russian were asked to supply the missing target word in each sentence. Cloze test scores are provided in Table 4. A two-way analysis of variance (ANOVA; with condition and context type as independent variables and cloze probability score as a dependent variable) was conducted for the Critical and Control conditions. The condition by context interaction as well as the main effects were not significant (condition by context: $F(2, 114) = 0.649$, $p = 0.52$, $\eta^2 = 0.011$; condition: $F(1, 114) = 0.287$, $p = 0.59$, $\eta^2 = 0.002$; context: $F(2, 114) = 1.025$, $p = 0.362$, $\eta^2 = 0.018$).

The sentences in all conditions were structurally simple in order not to create an additional processing difficulty. Due to the materials design constraints, frequency of the target words could not be ideally balanced across different conditions, but it was accounted for statistically during data analysis. Log

**Table 4** Mean log surface frequency and mean cloze test probability for target words (with standard deviations)

| Condition | Critical Frequency | Cloze | Control Frequency | Cloze | Filler Frequency | Cloze |
|---|---|---|---|---|---|---|
| Semantic | 0.92 | 65 | 1.68 | 77 | 1.25 | 67.83 |
| | (0.72) | (35.17) | (0.55) | (21.79) | (0.65) | (35.75) |
| Syntactic | 1.32 | 63 | 1.39 | 62 | | |
| | (1.04) | (34.35) | (0.61) | (30.71) | | |
| Morphological | 1.35 | 72 | 2.12 | 70 | | |
| | (0.67) | (27.07) | (0.38) | (32.77) | | |

surface frequency rather than log lemma frequency was calculated using the Russian national corpus (http://www.ruscorpora.ru/) in order to discriminate the items in the Morphological condition, where lemma frequency for any given minimal pair (e.g., "to cook" "cooks") would have yielded the same value (Table 4).

## Results

### AX Discrimination

The signal detection theory framework was adopted for the analysis in order to take into account participants' response bias commonly observed in discrimination tasks. Thus, participants' accuracy scores were calculated and converted to d́ scores. Correct responses to the "same" trials were coded as "hits," incorrect responses to the "different" trials were coded as "false alarms." The obtained d́ scores are plotted in Figure 1. A three-factor mixed-design ANOVA with consonant contrast type (two levels: Critical or Control) and vowel environment (two levels: /a/ or /i/) as the within-subjects factors, and language group (two levels: L1 or L2) as the between-subjects factor yielded a significant interaction between language group and consonant contrast type ($F(1, 174) = 22.32, p < 0.001, \eta^2 = 0.08$), a significant main effect of consonant contrast type ($F(1, 174) = 88.41, p < 0.001, \eta^2 = 0.12$), a significant main effect of vowel environment ($F(1, 216) = 37.87, p < 0.01, \eta^2 = 0.02$), and a significant between-subjects main effect of language group ($F(1, 42) = 136.8, p < 0.001, \eta^2 = 0.2$). Post hoc Tukey HSD tests indicated that L2 listeners' mean d́ scores in both the Control ($M = 2.06, SE = 0.05$) and the Critical ($M = 0.96, SE = 0.065$) conditions were significantly lower than those of L1 listeners in the same

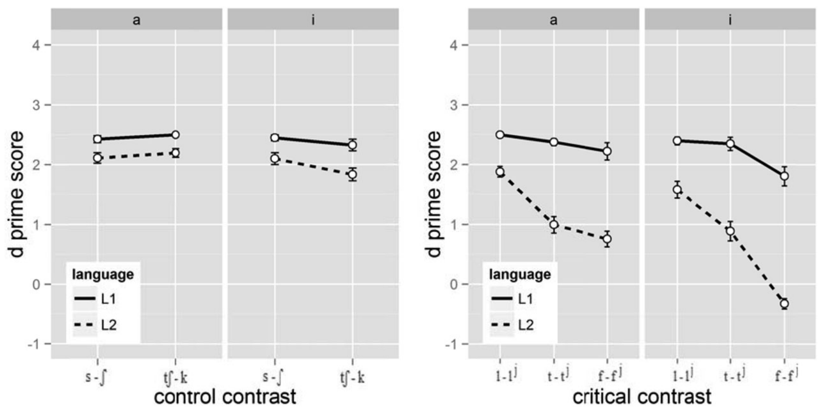


**Figure 1** Participants' d́ scores in the AX discrimination task.

conditions (Control: $M = 2.42$, $SE = 0.034$, Critical: $M = 2.27$, $SE = 0.047$) ($p < 0.05$ and $p < 0.001$, respectively). Importantly, the difference between Control and Critical conditions was significant only in the L2 group ($p < 0.001$), that is, their performance was on average significantly poorer in the Critical condition than in the Control condition. Post hoc comparisons also showed that while the vowel environment did not significantly affect L1 listeners' d́ scores, L2 listeners had more difficulty discriminating consonant contrasts in the /i/-environment ($M = 1.33$, $SE = 0.1$) than in the /a/-environment ($M = 1.68$, $SE = 0.08$) ($p < 0.001$).

A separate, two-factor ANOVA with language group as a between-subjects factor (two levels: L1 and L2) and consonant contrast (three levels: /l-l$^{j/}$, /t-t$^{j/}$, /f-f$^j$/) as a within-subjects factor was also carried out on the Critical condition data in order to examine the discrimination of individual target contrasts. A significant interaction between language group and consonant contrast ($F(2, 84) = 11.7$, $p < 0.001$, $\eta^2 = 0.08$), as well as a significant main effect of consonant contrast ($F(2, 84) = 73.01$, $p < 0.001$, $\eta^2 = 0.2$) and language ($F(1, 42) = 122.8$, $p < 0.001$, $\eta^2 = 0.41$) was found. Post hoc analyses revealed that while mean d́ scores were not significantly affected by the type of the consonant contrast in L1 group, they differed significantly in the L2 group. Thus, the /l-l$^j$/ contrast ($M = 1.73$, $SE = 0.09$) was on average easier than the /t–t$^j$/ contrast ($M = 0.94$, $SE = 0.136$), and the /f–f$^j$/ contrast ($M = 0.21$, $SE = 0.075$) was found to be perceptually the most difficult ($p < 0.001$).

### Listening Comprehension Task With Word Identification

First, in order to ensure that the participants were attending to meaning, their comprehension of the sentences was checked by analyzing the accuracy of responses to the 42 comprehension questions. All participants in both groups, L1 and L2, had high response accuracy scores (about 96% and 89%, respectively). Therefore, all participants were included in the subsequent data analysis, which proceeded in several steps: (a) overall error rate analysis of the word identification performance across all test conditions, (b) overall reaction time analysis, and (c) a separate analysis of the Critical condition.

#### *Error Rate Analysis*

Subjects' responses were scored as erroneous (error = 1) if they chose the wrong word between the two target words, and correct (error = 0) if they chose the correct word that occurred in the sentence. Similarly to L1 listeners, L2 listeners made few errors in the Filler and Control conditions for both congruent and incongruent sentences, but their error rate dramatically increased in the Critical condition, especially for incongruent items (Figure 2), suggesting that a degraded perceptual sensitivity to the hard/soft consonant contrasts presented a problem for correct word identification for the L2 but not the L1 listeners.

A generalized linear (logistic) mixed-effects model (GLMM) was performed on the data using the lme4 package (Bates & Maechler, 2010) in R statistical computing software (R Core Team, 2013) with a binomial family link function. The GLMM analysis was chosen over traditional ANOVA because it can account for possible individual differences among the participants and the variation that may exist in the stimulus materials. It also allows performing by-subject ($F_1$) and by-item ($F_2$) analyses within a single analytic framework. Condition (two levels: Critical and Control), congruency (two levels: Congruent and Incongruent), and language group of the participants (two levels: L1 and L2) were entered as fixed effects while subjects and items were entered as nested random effects with random intercepts. Filler condition was not included in the analysis because it differed from the Critical and Control conditions along many parameters. The best-fitting regression model included all main effects as well as three two-way interactions (condition by language, congruency by language, and condition by congruency) and one three-way interaction (condition by language by congruency). Table 5 presents the model's estimated coefficients for each predictor, their standard errors, the $z$ statistic (Wald test), and the associated $p$ values. The intercept (reference comparison) estimated the log odds of word identification error for the congruent items in the Control condition for the L1 group. Language (L2) parameter estimated the

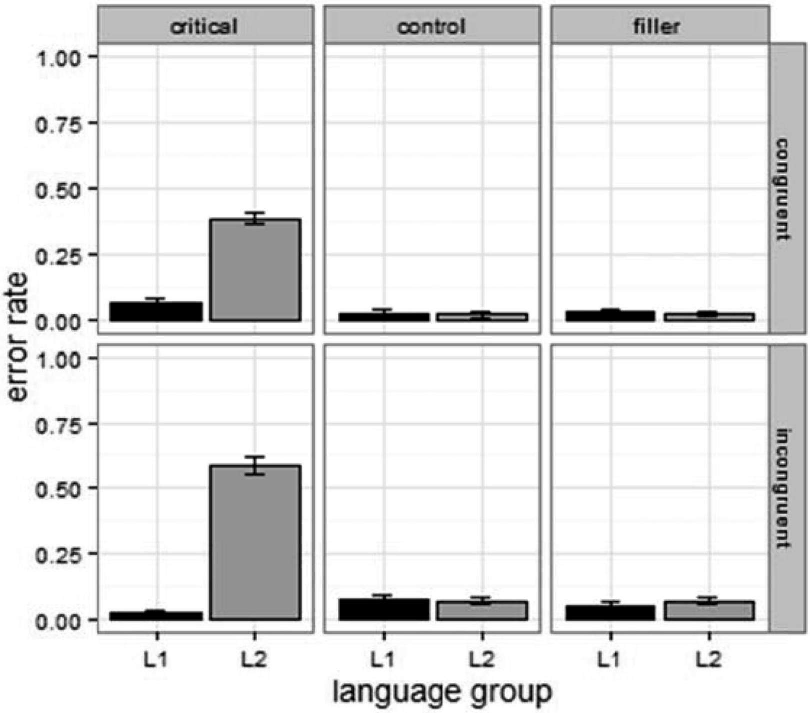


**Figure 2** Participants' mean error rate across all conditions in the listening comprehension task with word identification.

difference in error rate between the L1 and the L2 groups for corresponding conditions (in logits). The *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. The logistic regression coefficients are interpreted as the change in the log odds of the outcome brought by the change of the factor from one level to another (e.g., from L1 to L2, or from Congruent to Incongruent condition, etc.). For example, for the L1 listeners, a change from Control condition to Critical condition yields a change in log odds of error of 1.06.

The results of the GLMM demonstrate that L1 participants' error rates in the Critical and Control conditions were not statistically different from each other. L2 participants' performance was statistically different only in the Critical condition ($z = 3.56, p < 0.001$). Interestingly, the interaction between condition and congruency in the L1 group turned out to be significant ($z = -2.41, p < 0.05$).

**Table 5** Estimated coefficients from a mixed-effects model for participants' error rate in the listening comprehension task with word identification

| Fixed effects | Estimate | *SE* | *z* value |
|---|---|---|---|
| (Intercept) | **−3.82** | **0.54** | **−7.06**$^{***}$ |
| Critical | 1.06 | 0.62 | 1.7 |
| Incongruent | 1.05 | 0.60 | 1.75 |
| L2 | −0.29 | 0.65 | −0.44 |
| L2 : Critical | **2.56** | **0.72** | **3.56**$^{***}$ |
| L2 : Incongruent | 0.38 | 0.72 | 0.53 |
| Critical : Incongruent | **−1.98** | **0.82** | **−2.41**$^{*}$ |
| L2 : Critical : Incongruent | 1.41 | 0.92 | 1.53 |
| Random effects | Variance | *SD* | |
| Item | 0.25 | 0.50 | |
| Subject | 0.08 | 0.29 | |

*Note.* Bold indicates coefficients that are statistically significant, $^{***}p < 0.001$, $^{*}p < 0.05$.

*Reaction Time Analysis*

Reaction times were measured in milliseconds from the onset of the presentation of the two target words on the computer screen; therefore, they cannot reflect real-time costs of integrating target words in the sentential context; rather, they indicate a relative ease or difficulty of test-takers' decision making at the postprocessing stage and should be considered with caution. The results show that L2 participants' reaction times were systematically greater than those of L1 participants across all conditions (Figure 3). Their reaction times also increased considerably in the Critical condition relative to other conditions for both congruent and incongruent sentences.

Similarly to the error rate analysis, a GLMM was carried out with the participants' reaction time data as an outcome variable. The model compared response latencies only in the Critical and Control conditions with the exclusion of the Filler condition. Condition (two levels: Critical and Control), congruency (two levels: Congruent and Incongruent), and language group of the participants (two levels: L1 and L2) were entered as fixed effects while subjects and items were entered as nested random effects with random intercepts. The intercept estimated reaction times for the congruent items in the Control condition for the L1 group. The coefficients are interpreted as the change in the reaction times brought about by the change of a factor from one level to another. The estimated coefficients, standard errors, the *t* statistic and the associated *p* values

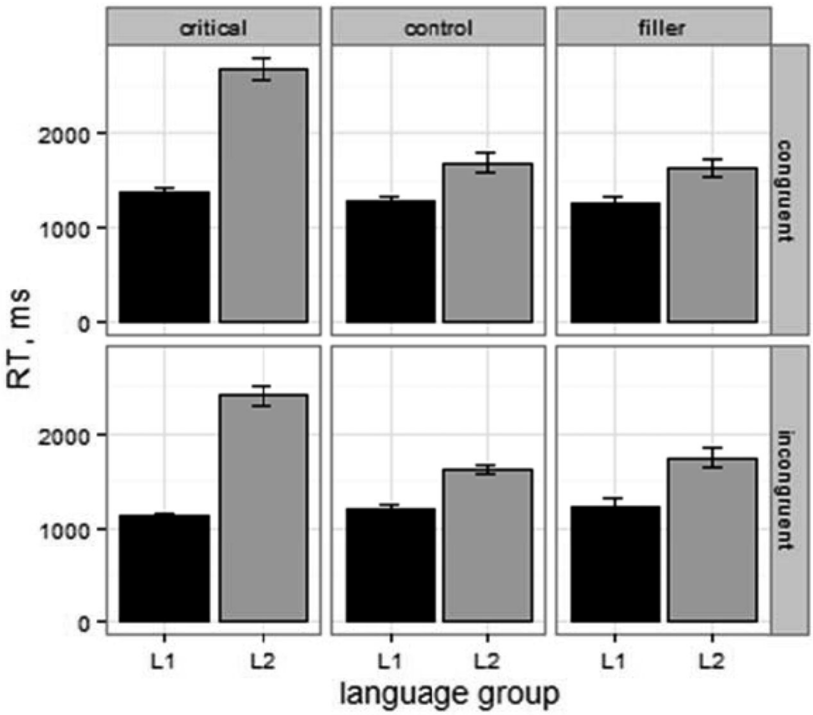


**Figure 3** Participants' mean reaction times (in milliseconds) across all conditions in the listening comprehension task with word identification.

are provided in Table 6. The results show that for the L1 group, reaction times in the Critical and Control conditions were not significantly different from each other. L2 speakers were 382.88 ms slower than L1 speakers in the Control condition, but they were statistically slower only in the Critical condition ($t = 6.32$, $SE = 149.02$, $p < 0.05$).

*Examining Contextual Bias Effect*

Overall error rate and reaction time analyses demonstrated that L2 listeners' error rate and reaction time latencies increased significantly in the Critical condition compared to other conditions. We analyzed this condition separately to examine whether particular context effects (Semantic, Morphological, or Syntactic) contributed to the increased error rate.[1]

First, L2 participants' knowledge of the target words (as indicated by their translation accuracy) in the Critical condition was checked. The responses

**Table 6** Estimated coefficients from a mixed-effects model for participants' reaction time in the listening comprehension task with word identification

| Fixed effects | Estimate | *SE* | *t* value |
|---|---|---|---|
| (Intercept) | **1,286.66** | **179.45** | **7.17*** |
| Critical | 80.15 | 135.63 | 0.591 |
| Incongruent | −97.31 | 124.8 | −0.78 |
| L2 | 382.88 | 206.53 | 1.854 |
| L2 : Critical | **941.62** | **149.02** | **6.32*** |
| L2 : Incongruent | 67.53 | 144.75 | 0.467 |
| Critical : Incongruent | −144.12 | 173.77 | −0.829 |
| L2 : Critical : Incongruent | −107.21 | 202.63 | −0.529 |
| Random effects | Variance | *SD* | |
| Item | 31,973 | 178.81 | |
| Subject | 266,369 | 516.11 | |

*Note.* $t$ value = Coefficient/SE, with $t$ values over 2.0 indicating that the coefficient is significantly different from zero (Gelman & Hill, 2007). Bold indicates coefficients that are statistically significant, $^{**}p < 0.05$.

were scored as correct if they were translations of any of the meanings of a polysemic item. For example, for ***мат***, both 'profanity' and 'stalemate' were considered acceptable translations. The total number of erroneous translations was counted for each target word. Twelve words (7 in the Semantic, 5 in the Syntactic condition) that received a high error rate (were incorrectly translated by more than 40% of participants) were removed. This resulted in the loss of 196 data points out of 960 (20.41% of the data). The mean error rate and reaction times as well as differences in error rate and reaction times between Congruent and Incongruent conditions for both the full dataset and the truncated dataset are presented in Table 7.

A GLMM analysis was performed on both the full dataset and the dataset with familiar words only. The type of contextual constraint (three levels: Semantic, Syntactic, and Morphological) and congruency (two levels: Congruent and Incongruent) as well as their interaction term were included as fixed effects. In order to control for perceptual salience, the type of the consonant contrast (three levels: /t-tʲ/, /l-lʲ/, and /f-fʲ/) was also included in the analysis. Additionally, because word frequency was not balanced across the three context conditions during the design of the experiment, we carried out a three-step mediation analysis (Baron & Kenny, 1986) to see whether word frequency mediated the relationship between participants' error rate and the type of the

**Table 7** L2 participants' mean error rate and reaction times in the Critical condition for ALL target words included and for only FAMILIAR words included in the analysis (with standard errors)

| Outcome variable | Congruency | Semantic | | Morphological | | Syntactic | |
|---|---|---|---|---|---|---|---|
| | | All | Familiar | All | Familiar | All | Familiar |
| Error rate | Congruent | 0.45 | 0.43 | 0.46 | 0.46 | 0.25 | 0.27 |
| | | (0.043) | (0.042) | (0.043) | (0.043) | (0.027) | (0.034) |
| | Incongruent | 0.42 | 0.42 | 0.76 | 0.76 | 0.58 | 0.67 |
| | | (0.05) | (0.064) | (0.046) | (0.046) | (0.048) | (0.06) |
| Error rate difference | | −0.03 | −0.01 | 0.30 | 0.30 | 0.33 | 0.40 |
| RT, ms | Congruent | 2670 | 2600 | 2955 | 2955 | 2448 | 2468 |
| | | (167) | (193.88) | (213.73) | (213.73) | (170.44) | (238.73) |
| | Incongruent | 2146 | 2139 | 2767 | 2767 | 2316 | 2347 |
| | | (157) | (182.67) | (222.25) | (222.25) | (155.87) | (163.57) |
| RT difference | | −524 | −461 | −188 | −188 | −132 | −121 |

**Table 8** Estimated coefficients from a mixed-effects model for participants' error rate in the Critical condition for ALL target words and for only FAMILIAR words included in the analysis

| Fixed effects | Estimate | | *SE* | | *z* value | |
|---|---|---|---|---|---|---|
| | All | Familiar | All | Familiar | All | Familiar |
| (Intercept) | 0.296 | −0.056 | 0.381 | 0.615 | 0.777 | −0.091 |
| Syntactic | **−0.885** | **−0.717** | **0.255** | **0.342** | **−3.47****| **−2.095*** |
| Morphological | −0.073 | 0.119 | 0.242 | 0.323 | −0.301 | 0.368 |
| Incongruent | −0.106 | 0.079 | 0.229 | 0.282 | −0.462 | 0.278 |
| Frequency | 0.0003 | 0.0002 | 0.0002 | 0.0003 | 1.664 | 0.721 |
| /l-l$^j$/ | **−0.783** | −0.366 | **0.391** | 0.641 | **−2.001*** | −0.570 |
| /t-t$^j$/ | −0.405 | −0.228 | 0.386 | 0.635 | −1.049 | −0.359 |
| Syntactic : Incongruent | **1.565** | **1.710** | **0.337** | **0.402** | **4.647**** | **4.249**** |
| Morphological : Incongruent | **1.423** | **1.204** | **0.335** | **0.373** | **4.245**** | **3.226**** |
| Random effects | Variance | | *SD* | | | |
| | All | Familiar | All | Familiar | | |
| Item | 0.035 | 0.101 | 0.187 | 0.317 | | |
| Subject | 0.117 | 0.111 | 0.342 | 0.333 | | |

*Note.* Bold indicates coefficients that are significantly different from zero, $^{**}p < 0.001$, $^{*}p < 0.01$.

contextual constraint. However, this mediation was not significant. Although frequency did not mediate the relationship between the dependent and the independent variable, it was included in the final model as a covariate in order to statistically adjust the dependent variable (error rate). The intercept estimated error response to the words in the congruent sentences of the Semantic condition. The logistic regression coefficients are interpreted as the change in the log odds of the outcome brought about by the change of the factor from one level to another (e.g., from Congruent to Incongruent condition, and so on). The estimated coefficients, standard errors, the $z$ statistic and the associated $p$ values are presented in Table 8.

The results of the GLMM indicated that L2 participants made significantly fewer mistakes on the congruent items in the Syntactic condition (full dataset: $z = -3.47, p < 0.001$; truncated dataset: $z = -2.095, p < 0.01$) compared to the Morphological and Semantic conditions. Notably, while L2 participants' error rate in the Semantic condition did not differ significantly between congruent and incongruent sentences (full dataset: $z = -0.462$, $p = 0.64$; truncated

**Table 9** L2 participants' mean error rate (with standard errors) broken down by sound type and congruency for ALL target words included in the analysis

| Congruency | Contrast type | Contextual constraint | | |
|---|---|---|---|---|
| | | Semantic | Syntactic | Morphological |
| Congruent | /f-fʲ/ | 0.38 (0.125) | - | - |
| Incongruent | /f-fʲ/ | 0.75 (0.11) | - | - |
| Congruent | /l-lʲ/ | 0.39 (0.06) | 0.23 (0.05) | - |
| Incongruent | /l-lʲ/ | 0.30 (0.07) | 0.55 (0.07) | - |
| Congruent | /t-tʲ/ | 0.51 (0.06) | 0.23 (0.04) | 0.46 (0.04) |
| Incongruent | /t-tʲ/ | 0.46 (0.05) | 0.63 (0.07) | 0.76 (0.05) |

dataset: $z = 0.278, p = 0.78$), error rate was significantly higher in incongruent sentences in the Morphological (full dataset: $z = 4.245$, $p < 0.001$; truncated dataset: $z = 3.226$, $p < 0.001$) and the Syntactic (full dataset: $z = 4.647$, $p < 0.001$; truncated dataset: $z = 4.249$, $p < 0.001$) conditions, suggesting that L2 participants experienced a stronger context bias in those two conditions. Indeed, while error rate in the Semantic condition was almost the same in the congruent and incongruent sentences, in the Syntactic and Morphological conditions it was about 30% larger in the incongruent sentences. Context bias effect increased in the Syntactic condition to 40% when unfamiliar words were excluded from the analysis.

With regard to the perceptual saliency of the target contrasts, the results revealed that the discrimination of the minimal pair with the labial obstruents /f-fʲ/ contrast (in the Semantic condition) was the hardest. Discrimination of the words with the coronal obstruents /t-tʲ/ and liquids /l-lʲ/ contrasts was easier, but only the words with the /l-lʲ/ contrast were identified significantly more accurately when all words were included in the analysis (full dataset: $z = -2.001$, $p < 0.05$). Because the number of the words with the three contrasts was not distributed equally across the three conditions (see Table 9), we cannot examine directly how the effect of perceptual saliency differed with the change of the type of the contextual constraint. Instead, we can examine how perceptual saliency affected word identification performance within the same context condition. To this end, separate comparisons for the Semantic and the Syntactic conditions were carried out (the Morphological condition was not included in the analysis because it contained words distinguished on the basis of the /t-tʲ/ contrast only). It was found that while there was no significant difference in overall error responses to words with either a /t-tʲ/ or an /l-lʲ/ contrast in the Syntactic condition, words with the /l-lʲ/ contrast in the Semantic condition were identified significantly more accurately ($z = 0.281$, $p < 0.05$). Such a pattern seems to parallel the results of the AX discrimination task, where it was observed that the /l-lʲ/ contrast presented a smaller discriminability difficulty for L2 listeners than the /t-tʲ/ and the /f-fʲ/ contrasts. Importantly, however, the type of the contrast did not affect the observed overall pattern of the results showing a context bias effect in the Morphological and the Syntactic conditions but not in the Semantic condition (note the pattern for the /t-tʲ/ contrast, which cuts through all three conditions; Table 9).

In light of these results, we examined whether L2 listeners' perceptual discrimination of the target phonological contrasts in the AX task (a task emphasizing low-level phonetic-acoustic properties of sounds) predicted their discrimination of the words distinguished on the basis of the same contrasts at the level of sentence processing. Three regression analyses were carried out such that the AX d́ score on each of the three critical contrasts (/l-lʲ/, /t-tʲ/ /f-fʲ/) was used to predict L2 listeners' mean error response to the words discriminated on the basis of the respective contrasts in the Critical condition of the listening task (e.g., d́ score on the /t–tʲ/ contrast was used to predict error rate to words with the same contrast, ***говорит*** ('speaks', /gɑvɑrʲit) – ***говорить*** ('to speak', /gɑvɑrʲitʲ/)). The regressions, however, did not come out significant for any of the contrast types.

## Discussion

Much of the work on L2 speech perception has focused on outlining the phonological difficulties that L2 speakers face in discrimination and categorization of L2 sounds. In the introductory part of this article, we explained how such phonological difficulties might cause lexical ambiguity and create a serious problem for L2 speech comprehension. The present study set out to examine whether L2 speakers can utilize information derived at higher levels of processing (semantic, morphological, and syntactic) to deal with such ambiguity at the sentence level in a way that is similar to L1 speakers (as evidenced by the literature on L1 ambiguity resolution). More importantly, the goal of the study was to explore which types of contextual cues affect L2 listeners' word identification the most.

First, we had to establish that our L2 participants—L1 English speakers of L2 Russian—experience difficulty with an L2 phonological contrast. The phonological contrast of consonantal hardness/softness was chosen as the target feature in this study, because it is known from previous studies that L2 Russian speakers show decreased perceptual sensitivity to this contrast, especially in the word-final position (Bondarko, 2005; Lukyanchenko & Gor, 2011). The results of the AX task support previous findings and demonstrate that the L2 participants in this study also showed a significantly lower discriminability of hard/soft consonants compared to the native Russian speakers, suggesting an approximate and unstable nature of the L2 phonological representations. Moreover, some hard/soft contrasts were found to present more difficulty than others. The /l-lʲ/ contrast was found to be the easiest, /t-tʲ/ occupied the intermediate position, and /f-fʲ/ presented the most difficulty. This indicates that the perceptual difficulty of the phonological feature of consonantal softness for L2 listeners is not instantiated equally for all consonants; rather, it depends on the properties of an individual consonant, such as place and manner of articulation (Kochetov, 2002).

Participants' perceptual difficulty with hard/soft consonants carried over to the listening comprehension task with word identification. For example, we observed that the L2 participants' unequal perceptual difficulty with the three tested contrasts carried over to their discrimination of the words that included the respective contrasts. In the Semantic condition, the words with the /l-lʲ/ contrast were discriminated more accurately compared to the /t-tʲ/ and /f-fʲ/ contrasts. This indicates that perceptually difficult L2 contrasts do not completely impede L2 word recognition: While they certainly create phonolexical ambiguity rendering word identification difficult, the confusable words may not be perceived as exact homophones (cf. Pallier et al., 2001; Weber & Cutler, 2004).

More importantly, the results of the listening comprehension task demonstrate that while both L1 and L2 listeners identified words with a high degree of accuracy in the Control condition, only L2 listeners demonstrated poor word identification performance in the Critical condition indicating fuzziness (but not homophony) of lexical representations, which are easily misinterpreted and miscoded. This difference in participants' performance means that when L2 listeners have the necessary phonological representations in place and can differentiate between the target phonological contrasts easily, they can rely on their bottom-up strategies to extract the necessary phonological information to guide them to the correct lexical decision even though information coming from higher-order levels of processing may be biasing them toward the wrong word (in the Incongruent condition). In contrast, when L2 speakers' phonological representations are fuzzy and unclear (Critical condition), they cannot fully trust the information extracted at the phonological level of processing because it is ambiguous. We hypothesized that, under such circumstances, L2 speakers would rely on contextual information to compensate for the phonolexical ambiguities as native speakers do. If this is true, they should demonstrate good performance in the Congruent condition (where both phonological and contextual cues direct them to the correct word) and poor performance in the Incongruent condition (where they are expected to be biased by the contextual information toward the wrong word). Quite remarkably, L2 listeners made many errors in the Critical condition in both congruent and incongruent items (Table 7). Such a pattern indicates that L2 listeners were not solely guided by the contextual cues in their word identification behavior in the case of phonolexical ambiguities, because if they were, their score in the Congruent condition would have been perfect and their error rate in the Incongruent condition would have been very high.

One explanation of the observed pattern could lie in the nature of the task used in this study. To remind the reader, the listeners had to reconcile conflicting cues in the Incongruent condition and choose the correct word despite context bias. In the course of the task, they became aware that solely relying on the context could lead them astray on a significant portion of the trials. While in the Control condition the L2 listeners were able to "override" the incongruent context and choose the right word on the basis of bottom-up information, phonological cues were not readily available in the Critical condition, which has resulted in some guessing behavior on congruent items. Importantly, however, had they adopted a guessing strategy by completely ignoring contextual cues, their accuracy would have been at chance on both congruent and incongruent items. Clearly, this is not the case for the Morphological and Syntactic conditions, where they made significantly more errors on the incongruent items than congruent items, both on familiar and less familiar words (Tables 7 and 9). In the Morphological and the Syntactic conditions, error rate increased by about 30% in the Incongruent condition relative to the Congruent one, suggesting a strong context bias effect. When unfamiliar words were excluded from the analysis, context bias in the Syntactic condition grew to 40%. This evidence indicates that L2 speakers do rely on context to distinguish between lexically ambiguous words, if not exhaustively, but at least to a great extent, and that morphosyntactic cues constrain word choice more effectively than semantic cues.

Such results seem to be at odds with the reviewed literature, where L2 speakers have been systematically shown to be more sensitive to semantic rather than morphosyntactic violations (e.g., Hahne, 2001). Several explanations can be offered to account for the observed results. One explanation of the weaker effect of the semantic constraints on L2 listeners' word identification could be due to the fact that semantic constraints are more specific while morphosyntactic constraints are more general (e.g., requiring a noun and not a verb but providing little information about its specific characteristics; Lee & Federmeier, 2009). Moreover, syntactic information is generally thought to be deterministic and definitive (and thus quite constraining) in a way that semantic information cannot be (Friederici, Pfeifer, & Hahne, 1993; see also Friederici, 2002). A sentence beginning with "Mary got soaked to the skin because she forgot the . . . " provides a semantically constraining context for the word "umbrella," but it cannot rule out other options like "raincoat." In contrast, the same sentence unambiguously and exhaustively specifies syntactic structure, that is, a noun phrase (e.g., "umbrella," "new umbrella," and so on) that should follow the determiner. In case of ambiguous information, semantic cues should therefore be less effective than syntactic cues in word disambiguation. This is exactly what Folk and Morris (2003) found. They did not observe ambiguity effects when ambiguity crossed syntactic categories (e.g., a park—to park), which suggests that syntactic category information becomes available first and mediates the semantic resolution process. With regard to L2 speakers, their lexicon is normally significantly smaller than that of an average adult native speaker; they have to operate on a reduced frequency range; and their lexical-semantic associations are weak and inefficient. All this prevents them from being able to firmly and rapidly commit to a single lexical entry even when the disambiguating context is made available (note, e.g., delayed context integration costs in lexical ambiguity resolution by L2 speakers in Elston-Güttler & Friederici, 2005, 2007). This is not to say that they are not building semantic predictions and expectations; rather, their predictions are of a less affirmative and definitive nature because the lexical processor is used to a higher degree of lexical uncertainty on a regular basis. For example, if they know only one meaning of the word "pen" ("an instrument for writing") and hear this word in a sentence "The pig stood in the pen," where "pen" means "an enclosure for domestic animals," they are likely to discover a new meaning of the word "pen." However, due to the less discerning and more tolerant lexical processor, they may similarly decide that the word "wug" in a sentence "The pig stood in the wug" is a real word and try to deduce a meaning. The task used in this study provided an offline measure of contextual effects (rather than moment-by-moment processes in context integration) where the test takers had to make a word identification decision after having heard the sentence in its entirety. Because no participant was exposed to both words from the same minimal pair and was never made aware of the underlying phonological juxtaposition during the task, L2 listeners could have incorrectly accepted context-incongruent words as fitting by mistakenly attributing some new meaning to them (similarly to the above example with "wug").

Similar findings were reported by Conrad (1983), who investigated whether nonnative listeners paid more attention to syntactic than semantic information in listening, compared to native speakers. She used a cloze test consisting of 55 sentences, in which every fifth word was removed and replaced by a blank of uniform length. The text was played to native and nonnative English listeners, who were then required to reconstruct the text by supplying words in the blanks on the answer sheet. Results indicated that native speakers of English used primarily semantic cues (i.e., information provided by the lexical-semantic relationships between words) to process aural texts, whereas both intermediate and advanced learners of English tended to direct their attention to syntactic cues (i.e., information provided by the grammatical structure of the sentences).

It is also possible that contextual cues operate differently for different classes of ambiguous words. In the present study, word categories of the target words differed across the three critical conditions. In the Semantic condition, we had noun-noun ambiguities; the Morphological condition included only verb-verb ambiguities; and in the Syntactic condition ambiguities crossed different syntactic categories (e.g., noun-verb). Studies examining processing distinctions between nouns and verbs have identified a significantly slower naming of verbs than nouns in the native and second languages (Faroqi-Shah & Waked, 2010; Szekely et al., 2005), dissociations of noun and verb retrieval in patients with aphasia (e.g., Caramazza & Hillis, 1991; Zingeser & Berndt, 1990), and different degrees of cortical activation of nouns and verbs (Yokoyama et al., 2006). Such noun-verb dissociation data are interpreted as evidence that lexical organization in the brain is governed by grammatical class (e.g., Caramazza & Hillis) with an implication that words within the same grammatical class compete for lexical selection more than words belonging to different grammatical categories (Dell, Oppenheim, & Kittredge, 2008; Levelt, Roelofs, & Meyer, 1999; Pechmann & Zerbst, 2002). Based on these assumptions, participants in the present study may have experienced more competition and uncertainty in the Semantic condition, which included ambiguities within the same grammatical class. For the same reason, the Syntactic condition was more effective in constraining word selection because it ruled out between-class competitors early in sentence comprehension. That is why minimal pairs like ***мат*** ('stalemate', /mɑt/) – ***мать*** ('mother', /mɑtʲ/) could have created more ambiguity than ***брат*** ('brother', /brɑt/) – ***брать*** ('to take', /brɑtʲ/). Now, in the Morphological condition, we also had ambiguities within the same grammatical class (verbs). However, this condition is different from the Semantic condition in that the phonological contrast marked the distinction between the two forms of the same verb rather than different verbs. It is quite possible that, besides being constrained by the grammatical class, word selection is also constrained by the inflectional paradigm, such that words belonging to different inflectional forms (e.g., a verb infinitive form versus a present tense form as in ***говорит*** (‘speaks’, /gɑvɑrʲit) – ***говорить*** (‘to speak’, /gɑvɑrʲitʲ/)) experience less competition, and, therefore, make morphological constraints more effective. Because the use of the verb infinitive in Russian is very straightforward and its usage is similar to that in English, L2 Russian speakers are able to benefit from this knowledge to constrain their word choice in sentence comprehension.

## Limitations and Conclusion

Although the present study provides some new insights into the problem of L2 phonolexical ambiguity at the sentence level, it also has certain unavoidable limitations associated with the difficulty of matching the target words along several parameters across conditions (e.g., word class, word frequency, perceptual saliency) and the choice of experimental methodology. The listening task used in the study asked the participants to identify which of the two words presented on the screen occurred in the sentence after they heard the sentence in its entirety. Because interpretation of the data is based on inferences from the measures taken at the endpoint of processing (i.e., a button press after the sentence presentation) rather than continuously, it is difficult to reveal the more dynamic aspects of L2 sentence processing under the given test conditions. Finer-grained, more implicit methods capable of capturing the ongoing process of listening behavior are needed in order to address the issue of *how* and *when* L2 listeners utilize different types of contextual information for phonolexical ambiguity resolution.

Despite these limitations, the preliminary findings of the present study suggest that unstable and unreliable perception of L2 phonological contrasts renders words discriminated on the basis of these contrasts ambiguous and confusing. In the face of such phonolexical ambiguity, L2 listeners utilize contextual information for meaning disambiguation, but with a different degree of success. Thus, morphosyntactic and syntactic constraints appear to be more effective in constraining the choice between two ambiguous words. These findings highlight the importance of including not only semantic and syntactic, but also grammatical and morphological levels of analysis in existing models of ambiguity resolution and speech comprehension. In the present study we have only examined a selected set of contextual constraints and morphosyntactic phenomena, and it remains to be seen whether the findings can be generalized to other kinds of contexts or to L2 speakers with other L1–L2 combinations and proficiency levels.

Final revised version accepted 13 December 2013

## Notes

1 The same GLMM model that was run with error rate as an outcome variable was also run with reaction time as an outcome variable. Except for the significant congruency effect ($t = -3.112$, $SE = 168.2$, $p < 0.05$), L2 participants' differences in reaction time latencies between Semantic, Morphological, and Syntactic conditions were not statistically significant. Therefore, the analysis of the reaction time data in the Critical condition was dropped from further discussion.

## References


Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychology research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

Bates, D., & Maechler, M. (2010). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375–37, http://CRAN.R-project.org/package=lme4

Bondarko, L. (2005). Phonetic and phonological aspects of the opposition of "soft" and "hard" consonants in the modern Russian language. *Speech Communication*, *47*, 7–14.

Bosch, L., Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology*, *12*, 189–222.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America*, *112*, 272–284.

Broersma, M. (2005). *Phonetic and lexical processing in a second language* (MPI Series in Psycholinguistics No.34) (Doctoral dissertation). Radboud University, Nijmegen, The Netherlands.

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System*, *36*, 22–34.

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *The Quarterly Journal of Experimental Psychology*, *64*, 74–95.

Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, *349*, 788–790.

Chambers, C., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1029–1040.

Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing: A reply to our commentators. *Applied Psycholinguistics*, *27*, 107–126.

Clahsen, H., & Felser, C. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42.

Conrad, L. (1983). Semantic versus syntactic cues in listening comprehension. *Studies in Second Language Acquisition*, *7*, 59–69.

Craig, C., Kim, B., Rhyner, P., & Chirillo, T. (1993). Effects of word predictability, child development and aging on time-gated speech recognition performance. *Journal of Speech and Hearing Research*, *36*, 832–841.

Cutler, A., Garcia Lecumberri, M. L., & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *Journal of the Acoustical Society of America*, *124*, 1264–1268.

Dell, G. S., Oppenheim, G. M., & Kittredge, A. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and Cognitive Processes*, *23*, 583–608.

Díaz, B., Mitterer, H., Broersma, M., & Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences*, *22*, 680–689.

Diehm, E. E. (1998). *Gestures and linguistic function in learning Russian: Production and perception studies of Russian palatalized consonants* (Doctoral dissertation). The Ohio State University, Columbus, OH.

Elston-Güttler, K., & Friederici, A. (2005). Native and L2 processing of homonyms in sentential context. *Journal of Memory and Language*, *52*, 256–283.

Elston-Güttler, K., & Friederici, A. (2007). Ambiguous words in sentences: Brain indices for native and non-native disambiguation. *Neuroscience Letters*, *414*, 85–89.

Faroqi-Shah, Y., & Waked, A. (2010). Grammatical category dissociation in multilingual aphasia. *Cognitive Neuropsychology*, *27*, 181–203.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*, 453–489.

Field, J. (2004). An insight into listeners' problems: Too much bottom-up or too much top-down? *System*, *32*, 363–377.

FitzPatrick, I., & Indefrey, P. (2010). Lexical competition in non-native speech comprehension. *Journal of Cognitive Neuroscience*, *22*, 1165–1178.

Folk, J., & Morris, R. (2003). Effects of syntactic category assignment on lexical ambiguity resolution in reading: An eye movement analysis. *Memory and Cognition*, *31*, 87–99.

Friederici, A. (2002). Towards a neural basis of auditory sentence processing. *TRENDS in Cognitive Sciences*, *6*, 78–84.

Friederici, A., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological, and syntactic violations. *Cognitive Brain Research*, *1*, 183–192.

Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Lexical ambiguity resolution and spoken word recognition: Bridging the gap. *Journal of Memory and Language*, *44*, 325–349.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Goldinger, S., Luce, P., Pisoni, D., & Marcario, J. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1211–1238.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r." *Neuropsychologi*a, *9*, 317–323.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*, 251–266.

Hahne, A., & Friederici, A. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*, 123–141.

Howie, S. (2001). Formant transitions of Russian palatalized and nonpalatalized syllables. *IULC Working Papers*, *1*, 1–22.

Hu, G., & Jiang, N. (2011). Semantic integration in listening comprehension in a second language: Evidence from cross-modal priming. In P. Trofimovich & K. McDonough (Eds.), *Applying priming research to L2 learning and teaching* (pp. 199–218). Amsterdam: John Benjamins.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Kochetov, A. (2002). *Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization*. New York: Routledge.

Lee, C., & Federmeier, K. (2009). Wave-ering: An ERP study of syntactic and semantic context effects on ambiguity resolution for noun/verb homographs. *Journal of Memory and Language*, *61*, 538–555.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral Brain Science*, *22*, 1–75.

Li, P., & Yip, M. (1998). Context effects and the processing of spoken homophones. *Reading and Writing*, *10*, 223–243.

Love, T., Maas, E., & Swinney, D. (2003). The influence of language exposure on lexical and syntactic language processing. *Experimental Psychology*, *50*, 204–216.

Lucas, M. (1999). Context effects in lexical access: A meta-analysis. *Memory and Cognition*, *27*, 385–398.

Lukyanchenko, A., & Gor, K. (2011). Perceptual correlates of phonological representations in heritage speakers and L2 learners. *Proceedings of the 35th*

*Annual Boston University Conference on Language Development* (pp. 414–426). Somerville, MA: Cascadilla Press.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*, 1–71.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

McClelland, J. L., Thomas, A., McCandliss, B. D., & Fiez, J. A. (1999). Understanding failures of learning: Hebbian learning, competition for representational space, and some preliminary experimental data. In J. Reggia, E. Ruppin, & D. Glanzman (Eds.), *Progress in brain research: Vol. 121. Disorders of brain, behavior, and cognition: The neurocomputational perspective* (pp. 75–80). Oxford, UK: Elsevier.

Mirman, D. (2008). Mechanisms of semantic ambiguity resolution: Insights from speech perception. *Research on Language and Computation*, *6*, 293–309.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., & Fujimura, O. (1981). An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception and Psychophysics*, *18*, 331–340.

Moss, H. E., & Marslen-Wilson, W. D. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1254–1276.

Navarra, J., Sebastián-Gallés, N., & Soto-Faraco, S. (2005). The perception of second language sounds in early bilinguals: New evidence from an implicit measure. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 912–918.

Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, *64*, B9–B17.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, *12*, 445–449.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, *24*, 501–528.

Pechmann, T., & Zerbst, D. (2002). The activation of word class information during speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 233–243.

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, http://www.R-project.org

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of cognitive neuroscience*, *18*, 2030–2048.

Salasoo, A., & Pisoni, D. (1985). Interaction of knowledge sources in auditory word identification. *Journal of Memory and Language*, *24*, 210–231.

Sanders, L., & Neville, H. (2003). An ERP study of continuous speech processing II. Segmentation, semantics, and syntax in non-native speakers. *Cognitive Brain Research*, *15*, 214–227.

Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, *52*, 240–255.

Sebastián-Gallés, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., & Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience*, *18*, 1277–1291.

Sebastián-Gallés, N., & Soto-Faraco, S. (1999). On-line processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, *72*, 111–123.

Strange, W., & Shafer, V. (2008). Speech perception in second learners: The re-education of selective perception. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153–191). Amsterdam: John Benjamins.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., et al. (2005). Times action and object naming. *Cortex*, *41*, 7–25.

Tabossi, P., & Zardon, F. (1993). Processing ambiguous words in context. *Journal of Memory and Language*, *32*, 359–372.

Tyler, L. K., & Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception and Psychophysics*, *34*, 409–420.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*, 1–25.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*, 231–256.

Yokoyama, S., Miyamoto, T., Riera, J., Kim, J., Akitsuki, Y., Iwata, K., et al. (2006). Cortical mechanisms involved in the processing of verbs: An fMRI study. *Journal of Cognitive Neuroscience*, *18*, 1304–1313.

Zingeser, L. B., & Berndt, R. S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language*, *39*, 14–32.

Zwitserlood, P. (1989). The locus of the effects of the sentential-semantic context in spoken word processing. *Cognition*, *32*, 25–64.

## Appendix

Stimulus materials used in the listening comprehension task with word identification.

| Pair | Context | Word #1 | Translation #1 | Word #2 | Translation #2 |
|---|---|---|---|---|---|
| **CRITICAL CONDITION** | | | | | |
| 1 | semantic | кров | shelter | кровь | blood |
| 2 | semantic | мат | obscenity | мать | mother |
| 3 | semantic | мел | chalk | мель | sandbank |
| 4 | semantic | угол | corner | уголь | coal |
| 5 | semantic | плот | raft | плоть | flesh |
| 6 | semantic | пыл | heat, ardor | пыль | dust |
| 7 | semantic | клад | treasure | кладь | load |
| 8 | semantic | гол | goal (in sports) | голь | the poor |
| 9 | semantic | суд | court | суть | gist, point |
| 10 | semantic | жест | gesture | жесть | tin |
| 1 | syntactic | брат | brother | брать | to take |
| 2 | syntactic | быт | lifestyle | быть | to be |
| 3 | syntactic | балет | ballet | болеть | to be sick |
| 4 | syntactic | билет | ticket | белеть | to turn white |
| 5 | syntactic | цел | safe, sound | цель | goal, aim |
| 6 | syntactic | был | was | быль | true fact |
| 7 | syntactic | дал | gave | даль | distance |
| 8 | syntactic | ел | ate | ель | fir/Christmas tree |
| 9 | syntactic | прибыл | arrived | прибыль | income |
| 10 | syntactic | стал | became | сталь | steel |
| 1 | morphological | говорит | speaks | говорить | to speak |
| 2 | morphological | готовит | cooks | готовить | to cook |
| 3 | morphological | ездит | goes, drives | ездить | to go, drive |
| 4 | morphological | чистит | cleans | чистить | to clean |
| 5 | morphological | звонит | calls | звонить | to call |
| 6 | morphological | помнит | remembers | помнить | to remember |
| 7 | morphological | ставит | puts | ставить | to put |
| 8 | morphological | стоит | costs | стоить | to cost |
| 9 | morphological | строит | builds | строить | to build |
| 10 | morphological | значит | means | значить | to mean |

| Pair | Context | Word #1 | Translation #1 | Word #2 | Translation #2 |
|---|---|---|---|---|---|
| **CONTROL CONDITION** | | | | | |
| 1 | semantic | дом | house | дочь | daughter |
| 2 | semantic | врач | doctor | враг | enemy |
| 3 | semantic | дед | grandfather | день | day |
| 4 | semantic | сын | son | сыр | cheese |
| 5 | semantic | сон | dream | сок | juice |
| 6 | semantic | нос | nose | ночь | night |
| 7 | semantic | рот | mouth | рок | rock |
| 8 | semantic | стих | poem | стиль | style |
| 9 | semantic | голос | voice | голод | hunger |
| 10 | semantic | бок | side | боль | pain |
| 1 | syntactic | успел | managed | успех | success |
| 2 | syntactic | брал | took | брак | marriage |
| 3 | syntactic | клал | put | класс | class |
| 4 | syntactic | шёл | walked | шок | shock |
| 5 | syntactic | жил | lived | жир | fat, grease |
| 6 | syntactic | мыл | washed | мышь | mouse |
| 7 | syntactic | пил | drank | пир | feast |
| 8 | syntactic | лёг | lay | лёд | ice |
| 9 | syntactic | мог | could | мощь | power |
| 10 | syntactic | рос | grew | роль | role |
| 1 | morphological | видел | saw | видеть | to see |
| 2 | morphological | играл | played | играть | to play |
| 3 | morphological | сказал | said | сказать | to say |
| 4 | morphological | думал | thought | думать | to think |
| 5 | morphological | сделал | did | сделать | to do |
| 6 | morphological | работал | worked | работать | to work |
| 7 | morphological | смотрел | looked | смотреть | to look |
| 8 | morphological | сидел | sat | сидеть | to sit |
| 9 | morphological | спросил | asked | спросить | to ask |
| 10 | morphological | любил | loved | любить | to love |
| **FILLER CONDITION** | | | | | |
| | filler | стол | table | стул | chair |
| 2 | filler | ложка | spoon | кошка | cat (she) |
| 3 | filler | чай | tea | май | May |
| 4 | filler | лак | nail polish | рак | crawfish, cancer |
| 5 | filler | рука | hand | мука | flour |

| Pair | Context | Word #1 | Translation #1 | Word #2 | Translation #2 |
|---|---|---|---|---|---|
| 6 | filler | друг | friend | круг | circle |
| 7 | filler | машина | car | малина | raspberry |
| 8 | filler | нога | leg | нора | hole, burrow |
| 9 | filler | салат | salad | халат | robe |
| 10 | filler | бал | ball | бар | bar |
| 11 | filler | забор | fence | собор | temple |
| 12 | filler | коза | goat | коса | braid, scythe |
| 13 | filler | пар | steam | дар | gift |
| 14 | filler | ваза | vase | виза | visa |
| 15 | filler | море | sea | горе | grief |
| 16 | filler | место | place | тесто | dough |
| 17 | filler | работа | work | забота | care |
| 18 | filler | банк | bank | танк | tank |
| 19 | filler | группа | group | груша | pear |
| 20 | filler | вид | view | гид | guide |
| 21 | filler | цена | price | жена | wife |
| 22 | filler | лес | forest | вес | weight |
| 23 | filler | вера | belief | вена | vein |
| 24 | filler | год | year | кот | cat (he) |
| 25 | filler | банан | banana | баран | sheep |
| 26 | filler | волна | wave | война | war |
| 27 | filler | мост | bridge | рост | height |
| 28 | filler | дача | summer cottage | дама | lady |
| 29 | filler | шапка | hat | папка | folder |
| 30 | filler | папа | dad | пара | couple |