Newcastle University

School of English Literature, Language and Linguistics

How do *ab-initio* second language learners start to detect words?

An exploratory study on Russian.

A thesis submitted to the Faculty of Humanities and Social Sciences of Newcastle University in fulfilment of the requirements for the degree of Doctor of Philosophy

Natalia Pavlovskaya

August 2019

Abstract

The speech stream is a continuum and discrete units, e.g. words, cannot be identified from the signal alone. How language learners *segment* (i.e. recognise and store words) in the speech stream has typically been explored with respect to children (e.g., Jusczyk *et al.* 1994; 1999a,b). Researchers have only recently begun to examine how adult second language learners segment an unfamiliar natural language after ‘first exposure’ without instruction (Gullberg *et al.* 2010, 2012; Carroll 2012, 2013, 2014; Shoemaker & Rast 2013).

I report on a study of how 28 English-speaking adults begin to segment words after hearing them in fluent Russian speech during four sessions. The study explored the following questions: (1) Does participant' ability to identify words increase over sessions? (2) Do participants rely on segmentation cues such as phonotactics, word-initial stress, and word length? (3) If so, how do these cues interact? (4) Can learners generalise to the novel examples? (5) Are there differences between linguistically trained and naïve participants?

Each day for four successive days, 28 participants were exposed to audio input in Russian for seven minutes (= 28 minutes exposure). Input comprised of 48 sentences of natural speech with target words embedded in a sentence medial position. After each exposure phase, participants were tested on their detection abilities of words they heard in the input as opposed to words they did not hear using three tasks: a word recognition task, a forced-choice task, and a cognate identification task. The word identification and the forced-choice tasks investigated if participants could detect words they heard in the input as opposed to words they had not heard. The purpose of the cognate identification task was to eliminate those participants who might not have been paying sufficient attention to the input (which was uncontrolled in the previous studies on first exposure).

A word recognition and a forced-choice task conducted each day showed that segmentation improved significantly over time. Segmentation patterns reflected the influence of English phonotactics, sensitivity to weak-strong stress, and the interaction of the two, which, particularly for the word recognition task, stems from participants subconscious analysis of Russian. Also, participants could generalise phonotactic patterns of Russian to novel words. The study did not find a difference between linguistically trained and naive participants. The study concludes that beyond native language bias, adults deploy the various segmentation mechanisms similar to those children use.

Acknowledgements

This work would not have existed without the help of many people. First and foremost, I would like to express my deepest gratitude to my supervisors Prof Martha Young-Scholten and Dr Nick Riches for inspiring me with this fascinating topic and for continuous support during my PhD study.

Besides my supervisors, I would like to thank my committee members Dr S.J. Hannahs and

Prof Leah Roberts for making my defence to be an enjoyable moment, and their insightful comments, questions and encouragement which allowed me to improve the thesis.

I am also sincerely thankful to Dr Alex Ho-Cheong Leung for the time he spared to help me at the beginning stage of my PhD. Moreover, I would like to thank all the staff from the School of English Literature, Language and Linguistics for being constantly available to lend a helping hand. In particular, this work has benefited from discussions with Dr S.J. Hannahs, Dr Geoffrey Poole, Prof Maggie Tallerman and Dr Danielle Turton, and of course, my supervisors. In addition, I am grateful to Dr Ghada Khattab and Dennis Bates from the School of Education, Communication and Language Sciences for their guidance with the use of phonetic equipment in a sound-proof booth when I needed to record experimental stimuli.

Importantly, I would like to thank Newcastle University, the School of English Literature, Language and Linguistics for the School Discretionary Award, which helped me to cover expenses related to my experiment. Furthermore, I acknowledge all of my 29 participants' tremendous help.

I am forever grateful for the wealth of moral support I received from my mom Tatiana, sisters Irina and Olga, brother-in-law John, and my husband Victor.

Table of contents

List of figures

List of tables

List of abbreviations

ˈ	Denotes word-stress in IPA transcriptions (stress on the vowel to the right)
*	Ungrammatical
/ /	Denotes phonemic representations
[ ]	Denotes phonetic representations
C	A consonant
CI	Confidence Intervals
IQR	Interquartile Range
L1	First Language
L2	Second Language
MDH	Markedness Differential Hypothesis
MSD	Minimal Sonority Distance
MSS	Metrical Segmentation Strategy
OMP	Onset Maximisation Principle
OR	Odds Ratio
SSP	Sonority Sequencing Principle
V	A vowel

Chapter 1. Introduction

1.1 The segmentation problem

The comprehension of speech requires the breaking up of continuous speech into discrete meaningful units also known as *words*. The process of breaking up continuous speech into words is called *segmentation*. This is a complicated process, for several reasons. Firstly, speech is continuous, so that when people speak, the words are not separated by explicit pauses (Cole & Jakimik 1980). This makes the recognition of words in a speech context more difficult than recognising them in isolation. Furthermore, many speech phonemes within words may change depending on their phonetic context. For instance, sounds may become more like the sounds they are adjacent to, such as when the sound /n/ (as in 'piano' [pɪˈanəʊ]) can turn into a velarised nasal [ŋ] when it is followed by a velar stop, as in 'language' [ˈlaŋgwɪdʒ]. This phonological process is called *coarticulation*. Moreover, there is a vast amount of variability among speakers in terms of how sounds are articulated. For example: on average, women's voices have higher pitch, while men's voices have lower pitch; there are differences in the voices of pre-puberty and post-puberty talkers; the rate of the speech stream may vary from very slow to very fast among speakers (Miller & Liberman 1979). Furthermore, many languages have a plenitude of local varieties.

We know with certainty that the above factors associated with the complexities of speech processing are important due to studies of spontaneous misperceptions of speech[1] (Cutler & Butterfield 1992; Bond 1999) and through observations made of speech recognition by machines (Bernstein & Franco 1996; Brent 1999). Machines are notoriously bad at speech recognition, mainly because they cannot merely rely on pauses to identify word boundaries.

[1] It is during the process of decoding speech that slips of the ear can occur (Bond 1999).

However, most speakers of their first (L1) or native language (NL)[2], regardless of the language concerned, find a way to extract linguistically meaningful units from the speech stream effortlessly. An extensive study which was conducted on L1 word detection abilities by infants indicates that 4-month-olds can recognise the sound patterns of their own names in continuous speech (Mandel *et al.* 1995), and that infants between 7.5- and 10.5-months of age can use allophonic, phonotactic, and even weak-strong stress cues for the identification of word boundaries (e.g., Friederici & Wessels 1993; Jusczyk *et al.* 1999 a, b). For example, 9-month old Dutch infants prefer to listen to isolated lists of words which conform to the phonotactics of Dutch such as 'bref' and 'murt' rather than those words which are disallowed in Dutch such as 'rtum' and 'febr'; and 10.5-month old native English infants can use aspiration as a cue to distinguish 'night rates' as opposed to 'nitrates', just as they can isolate weak-strong words such as 'guitar' from passages after they have heard those words in isolation.

The task of segmentation for second language (L2)[3] learners is more difficult, especially for post-puberty L2 learners, because they already know a language. Such learners are biased in their listening due to their L1, *via* language *transfer,* which means that learning one language affects the subsequent learning of another language (Lado 1957). Roughly speaking, knowledge of the L1 may result in advantages due to *positive transfer* when knowing an L1 facilitates L2 acquisition through the existence of similar structures, as well as disadvantages due *negative transfer* when knowing an L1 interferes with L2 acquisition because the structures of the L1 and L2 are different (in other words, when an L1 structure can be transferred into the L2, but it will be incorrect).

A possible strategy which learners may apply for the purpose of speech segmentation to rely on their knowledge of words which are learned in isolation. Although detecting words in a sequential

[2] A native language (L1) is a language which a person has been exposed to from birth. In this thesis, terms such as 'native language', 'first language' and 'mother tongue' are used interchangeably.

[3] Researchers who work within a *first exposure paradigm* (discussed in Chapter 3), when they talk about exposure to an unfamiliar language, call this language a 'target language'. However, throughout this thesis, whenever I talk about a non-native language or any language learned after the age of four in addition to the one learned from birth, it is called a second language (L2).

context after hearing them in isolation may work (for example, in English, Jusczyk & Aslin 1995), but in general, this strategy is problematic in L1 acquisition because words which occur in isolation are rare. Research on L1 speech segmentation has shown that utterances directed to infants contain more than one word (van de Weijer 1998), and when mothers were asked to teach their children new words, only 20% of the words were produced in isolation (Woodward & Aslin 1990). In addition, many words do not occur in isolation, such as the articles 'a' or 'an' and 'the' and many function words in English. On the contrary, in L2 speech segmentation, it is more likely that L2 learners are exposed to more isolated word forms than an infant due to the methods used in L2 teaching. For instance, words in isolation are frequent in typical instruction and, certainly, word boundaries are extremely clear in the ample written text to which classroom learners are commonly exposed. Additionally, it would be impossible for them to store every word of a language in a mental lexicon. Consequently, relying on knowledge of words alone for segmentation cannot explain how a continuous speech stream can come to be segmented by babies and also adults who encounter an unfamiliar L2 for the first time.

So, how does a language learner start to segment words? This question has typically been asked regarding an L1. Extensive research on L1 speech perception by infants has been conducted by Peter W. Jusczyk and his colleagues. In comparison, although nowadays we know more about L2 learners' segmentation abilities after adults have already accumulated some knowledge of the L2, segmentation abilities in L2 acquisition have received relatively less attention compared to in L1 learners. Only recently have researchers started to investigate how adult learners who encounter an L2 for the first time in naturalistic conditions (known as *ab-initio*[4] learners), and we know surprisingly little about such L2 learners segment speech. Chapter 3 reviews the relevant research. In general, an indication of findings from these studies is that novel words can be segmented and even mapped onto meaning from aurally presented unfamiliar input (e.g. Rast 2008, 2010;

[4] Also *first exposure* learners in work by Susanne Carroll and Rebekah Rast along with their colleagues, or *minimal exposure learners* in work by Marianne Gullberg, Leah Roberts and Christine Dimroth.

Gullberg *et al.* 2010, 2012; Carroll 2012, 2014; Shoemaker & Rast 2013). However, it is still not clear what role the L1 and speech segmentation cues which have been shown to influence L1 and L2 speech segmentation affect these *ab-initio* learners' segmentation abilities. The present thesis describes another study of *ab-initio* learners' abilities, but before formulating the research aim, more details on speech segmentation cues and L1 transfer need to be provided. The next section accomplishes this.

1.1.1 Speech cues

It has been suggested that language learners can use certain perceptual predispositions to extract words from speech. That is, there must be bits of information in a language which, in the absence of explicit signals, are indicative of the beginnings and endings of words. Researchers have proposed various cues which could potentially facilitate speech segmentation. Among them are lexico-semantic, syntactic, and phonological information. However, studies which have examined how syntactic and lexico-semantic cues could help learners with word identification are scarce (but see, for instance, Sanders & Neville 2000, 2002; Hanulikova *et al.* 2010, 2011). It is likely that knowledge about sentence structure and the meanings of words could aid language learners in extracting words. However, as discussed above in Section 1.1, this type of knowledge may be unhelpful until a learner has succeeded in identifying and extracting the acoustic forms of words (Sanders & Neville 2000: 1). Many types of phonological cues have been proposed which could potentially aid language learners in detecting the beginnings and endings of words. These include:

(1) knowledge about familiar sounds and words (Jusczyk & Aslin 1995; Newman *et al.* 2003; Tsay & Jusczyk 2003);

(2) allophonic variation can be a cue as it concerns the variability of how a phoneme can be pronounced depending on its position in a word (Hohne & Jusczyk 1994; Jusczyk *et al.* 1999a), and in L2 speech segmentation (Altenberg 2005a; Ito & Strange 2009);

(3) probabilistic cues or in other words, statistical information (Saffran *et al.* 1996a; Brent & Cartwright 1996; Saffran *et al.* 1996b; Aslin *et al.* 1998; Johnson & Jusczyk 2001);

(4) the prosodic structure of a language and numerous studies of English have shown that in L1 speech segmentation, bisyllabic words which are stressed on the first syllable are segmented better than those which are stressed on the second syllable (Cutler & Norris 1988; Cutler 1990, 1994; Cutler & Butterfield 1992; Jusczyk *et al.* 1993b; Turk *et al.* 1995; Jusczyk *et al.* 1999b), and in L2 speech segmentation several studies by Archibald (1992, 1993) and Hart (1998);

(5) phonetic and phonotactic constraints which pose restrictions on the possible ordering of phonemes may be used in L1 speech segmentation (Brown & Hildum 1956; Messer 1967; Friederici & Wessels 1993; Jusczyk *et al.* 1993a, 1994; Mattys & Jusczyk 2001); and in L2 speech segmentation (Altenberg & Cairns 1983; Weber 2000; Altenberg 2005b; Weber & Cutler 2006).

It should be clear from the list above that L1 and L2 learners could apply several different cues for segmentation. However, phonotactic and prosodic cues are the most widely studied cues and have been shown to influence the speech segmentation of both L1 and L2 learners. Additionally, it is conceivable that the extraction of words from the speech stream presupposes the integration of more than one cue. Research from several decades ago shows that L1 adults can integrate two cues (Vitevitch *et al.* 1997; McQueen 1998; Mattys *et al.* 1999). The effect of multiple cues for L2 speech segmentation has only recently started to be examined for cases of a lexico-semantic cue in combination with one or two phonological or syntactic cues (e.g. Sanders & Neville 2000, 2002; Hanulikova *et al.* 2010, 2011). However, no studies to my knowledge have looked at the integration of two phonological cues for speech segmentation by L2 learners. This is surprising because the combination of different cues is essential and in real-life word detection, we are likely to integrate cues. To the present author's knowledge, the current study is the first to explore the integration of cues in *ab-initio* learning, thereby addressing this significant gap in the literature and contributing to our knowledge of how cue integration can facilitate segmentation. The present study aims to fill this gap by investigating how an adult, without formal instruction, segments the speech stream of a completely unfamiliar natural language that is presented in the form of aural input, and whether or not L1 learners can use phonotactic, prosodic, and word length (measured in terms of numbers of syllables) within words as cues for detecting word boundaries.

1.1.2 Transfer

What also needs to be mentioned here is the central issue in L2 speech segmentation studies (and fact in any L2 study) of the role the L1 plays in L2 acquisition (see, e.g. Major 2008 for discussion). Nowadays researchers consistently agree that not all elements that are similar are easy to acquire (Flege 1992; Flege *et al.* 1995), and not all elements that are different are difficult to acquire, as was originally proposed by Lado (1957) in his formulation of the *Contrastive Analysis Hypothesis* (CAH). Not long after the introduction of the CAH, Corder (1967) suggested that systematicity of learners' errors and the source should be considered in order to understand learners' *interlanguage* (Selinker 1972). It has long been known that, to understand how the L1 influences the acquisition of a new phonological system, not only do phonological differences need to be considered but also the universals of phonology (see below).

Nearly every if not all of the studies discussed in Chapter 2 and Chapter 3 on how L2 learners segment speech by means of phonological cues show that learners are affected by their L1 in some form or another. Therefore, in the process of finding out whether or not L1 English *ab-initio* learners can detect words in their new language (Russian), the present study also aims to examine whether or not learners are guided by native language phonological constraints, such as phonotactics and trochaic prosodic patterns which operate in their L1 English, e.g. [ˈklʲevʲɪr] and [ˈblʲinʲɪk], or if these learners are sensitive to the structures which exist in Russian and are not available from the L1 (e.g. words starting with the following onset clusters: *xl-*, *kn-*, *tv-*, and *ʃk-*). More specifically, going back to the discussion of the role of phonological cues in speech segmentation, this study aims to investigate if *ab-initio* learners can rely on phonological cues such as phonotactic cues (native vs non-native phonotactics), prosodic cues (strong-weak vs weak-strong bisyllabic words) and length of words cues (monosyllabic vs bisyllabic words) for the speech segmentation of Russian. The next section provides a rationale for the choice of English and Russian as a source and target languages, respectively.

1.2 Motivation for English L1 and Russian L2 and inclusion of universals

English was chosen as a source language for practical reasons because the present study was conducted in Newcastle upon Tyne, United Kingdom (UK), and so it was convenient for the researcher to recruit L1 English speakers. As a matter of fact, the Russian language was also chosen for practical reasons, because it is the present researcher's native language. At the same time, Russian was chosen because it was anticipated that exposure to Russian is relatively uncommon in the UK, and this is likely to guarantee *ab-initio* learning. Furthermore, L1 English and L2 Russian represent an interesting pair of languages from a phonological point of view for an investigation of transfer. These reasons are elaborated on below.

Office for National Statistics published information from the census in 2011, which collected information about language within England and Wales (Potter-Collins 2013). According to this information, almost 92% of the population speak English as their first language, and the remaining 8% speak another language as their main language. The most widely spoken other language was Polish (with 1% of the population or nearly half a million people using it), followed by Indian languages (e.g. Panjabi, Urdu, Bengali and Gujarati, with 0.5% and 0.4% of population), Arabic, French and 'all other' Chinese[5] (with 0.3% of population), and other European and non-European languages, such as Portuguese, Romanian or Persian comprising no more than 0.2% and 0.1% of population using these languages). Russian is not listed. This is good evidence that exposure to

[5] All other Chinese excludes Mandarin Chinese and Cantonese Chinese.

Russian is uncommon in the UK. However, with Polish being the main language of 1% of the UK population, English speakers may have been exposed to another Slavic language.

As for the L1 English and L2 Russian language pair, how the two differ in interesting ways is discussed in detail in Chapter 4. Briefly, both languages belong to the Indo-European language family, but English is West Germanic and Russian is East Slavic. Their phonetic inventories are similar, with some segments present in English being absent in Russian and *vice versa.* English stress is complicated but generally morphophonologicaly predictable, while stress in Russian seems to be less predictable (see Chapter 4). As Russian allows stress to fall on any vowel within a stem, stress placement in the words used in the tasks for the present study was kept constant with bisyllabic words being stressed either on the first or second syllable. However, it was anticipated that English learners would be affected by the dominant strong-weak stress placement in English.

What is interesting about English and Russian is the nature of their phonotactic constraints from the point of view of *markedness*. This term was first coined by Trubetzkoy and Jakobson in the 1930s and has since been given various definitions, all of which refer to the likelihood of occurrence of a particular linguistic phenomenon (see Gurevich 2001 and Haspelmath 2006 for discussion). The present study, as would perhaps most L2 phonology acquisition studies, adopts the definition by Eckman (1977: 320), the formulation of which involves implicational hierarchies:

> A phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A.

For instance, in terms of branching onsets, Russian allows violations of the Sonority Sequencing Principle (SSP), for example, in the clusters *rt-*, *lʒ-, lg-,* and *ptʃʲ-* (see also Chapters 2 and 4). This means that sonority plateaus such as *kn-* and *tv-* and sonority rises like *hl-* and *sr-* are possible in Russian. English allows sonority rises, which means that it also allows CV syllables.

Sonority relates to the *Markedness Differential Hypothesis* (MDH), which predicts that a speaker of a more marked language will find it easier to learn an L2 which is less marked; if an L2 is more marked than an L1, then it is more difficult to acquire (Eckman 1977). So, in Eckman's sense, the relative degree of markedness corresponds to the relative degree of difficulty or in other words, marked = more complex, and unmarked = simpler[6]. When English and Russian are compared with respect to onset phonotactics, the English phonotactic structure is *less marked*, and is thus simpler than Russian, and consequently Russian is *more marked*, and has onset phonotactics which are more complex than those of English.

Many studies have shown that children, as well as L2 learners, are affected by markedness. In L2 phonology, Broselow and Finer (1991) showed that Hindi, Japanese and Korean learners of English tended to simplify English CR consonant clusters into unmarked CV syllables through the insertion of epenthetic vowels (e.g. [pəruf] instead of 'proof'), the deletion of a second consonant in a cluster (e.g. [puf] instead of 'proof'), or the replacement of one of the consonants by another with a different manner of articulation known as metathesis (e.g. [pjuf] instead of 'proof'). Anderson (1987) also examined the acquisition of English onsets and codas, by speakers of Egyptian Arabic, Mandarin Chinese and Amoy Chinese. The results were consistent with the MDH in that Arabic-speaking subjects were most accurate, as their L1 allows clusters, whereas Chinese does not. A study by Carlisle (1991) showed that L2 learners make more errors in the production of more marked syllables than less marked ones in terms of sonority distance; that is, obstruent + nasal clusters were modified more often than obstruent + liquid clusters. Another example is a study of some relevance to the present thesis. Ostapenko (2005) examined the production of Russian clusters of L1 English learners at a high-proficiency level. Just like in other studies cited above, she found support for the MDH in how her learners applied several onset

[6] Also see Rice's (1999) comprehensive list of definitions of markedness in the literature. In addition to marked = more complex, there are also such notions such as being less natural, more specific, appearing in fewer grammars, later in language acquisition, harder to articulate; and unmarked = simpler, more natural, more general, appearing in more grammars, earlier in language acquisition, and easier to articulate.

simplification strategies such as epenthesis, deletion, and consonant substitution in their attempts to pronounce Russian words with onsets which violated the SSP; for example, [stikatʲ] or [staktʲ] instead of [stkatʲ], and [patˈkrʲeɪtʲ] instead of [patˈklʲeɪtʲ].

Finally, from the point of view of the onsets and offsets (beginnings and endings of words), numerous studies have shown that, although language learners are sensitive to both in word detection, it is generally assumed that onsets are more salient; for instance as indicated in studies by Messer (1967), McQueen (1998) and Mattys and Jusczyk (2001). Additionally, the *Onset Maximisation Principle* (OMP) states that languages tend to maximise the beginnings of syllables – the onsets and to minimise syllable endings – the codas (Selkirk 1984). This means that when one encounters an ambiguous parsing of a sequence of sounds, for instance, a Russian word for 'to build' which is *postroit* [paˈstroɪtʲ], how does one parse it? Is it *always-stroit*? Or *pos-troit*? Or perhaps *post-roit*? Relying on the OMP and Russian phonotactics, it is most likely that the former two parsings are the most natural as Russian allows both tri-consonantal onsets such as *str-*, and biconsonantal onsets such as *tr-*. The present study is concerned with the effects of word onsets on the ability to detect words from novel continuous speech in Russian. Additionally, to avoid confusion associated with three-member onsets, the present study focuses on the effect of biconsonantal branching onsets in Russian.

1.3 Choosing *ab-initio* study participants

In psycholinguistic studies such as the one carried out for the present thesis, it is common to use a sample of undergraduate students. However, not everyone agrees that the results obtained from these studies can be generalised to a broader population (e.g. Henrich *et al.* 2010). There has been debate among researchers as to which types of participants are best for testing linguistic hypotheses such as those in the present thesis. For instance, it has been thought that, in general, for testing semantic and syntactic hypotheses, the best subjects are those who have some form of linguistic training since these people are somewhat aware of language as an object of investigation, unlike linguistically naïve participants (Edelman & Christiansen 2003). However, Gibson and Fedorenko (2010: 98) question the usefulness of those types of knowledge which are present in theoretically aware participants but not in those who are theoretically naïve. They attribute the better ability of linguistically aware participants to the presence of *cognitive biases*; that is, to a systematically divergent pattern of behaviour from what is considered rational or normal in a cognitive task (Tversky & Kahneman 1974; also see Haselton *et al.* 2005). For instance, linguistically aware participants may unintentionally guess a researcher's hypotheses and apply their metalinguistic knowledge when making judgements in linguistics experiments. Therefore, Gibson and Fedorenko (2010), as well as Arunachalam (2013), suggest that a linguistically naïve sample should be preferred to a sample of linguists with language training. In studies on the effect of phonological cues in either an L1 or L2, or any study on *ab-initio* learners, linguistically aware participants are hardly ever used, perhaps with the exception of studies by Rebekah Rast and her colleagues which utilised a sample of L1 French subjects with knowledge of various other L2s and L3s who were taking a training programme to become French foreign language teachers. Participants in their experiments could have been biased due to their *metalinguistic knowledge* of language. Metalinguistic knowledge[7] commonly refers to language learners' cognitive ability to analyse language explicitly (e.g. Bialystok 1979; Elder *et al.* 1999; R. Ellis 2004). Elder (2009: 137) suggests that metalinguistic knowledge is learned *via* formal instruction, is not intuitive but consciously controlled, and is not automatic and therefore is difficult to access during spontaneous language production. The formal learning of a second language in an instructional context is most likely to contribute to the development of metalinguistic knowledge (Bialystok 1991 and Elder 2009). For example, intensive instruction in grammar was shown to be responsible for the high levels of metalinguistic knowledge among L1 learners of Chinese at the initial stage of learning when their results were compared with those

[7] Some researchers distinguish between metalinguistic knowledge and metalinguistic awareness (e.g. Masny 1987; Elder 2009); however, in the present thesis these terms are used interchangeably.

who received meaning-focused instruction (Elder & Manwaring 2004). This is generally true, though with the possible exception, of a study by Brown and Hildum (1956) who, in addition to a sample of students of psychology, utilised a sample of students who had studied descriptive linguistics. Both groups were exposed to real English words, including phonotactically legal and phonotactically illegal non-words, under conditions of noise and were required to write those words down as they heard them. The linguistically sophisticated group were told that some words which they would hear would be illegal. Hypothetically, they could have relied on their metalinguistic knowledge. However, the results showed that linguistically sophisticated participants performed in the same way as naïve participants by transcribing real words and legal non-words best. This means that, when it comes to phonotactic judgements, even when linguistically sophisticated participants were explicitly instructed to expect illegal sequences, were not any better than linguistically naïve participants. This study also pointed to the robustness of the effect of phonotactic constraints. This is important since these findings contradict claims that for testing a linguistic hypothesis a naïve sample should be used. Consequently, the present study, uniquely among studies of *ab-initio* learners, introduces linguistic or language knowledge background as a group variable where all participants who participated were divided into linguistically sophisticated and linguistically naïve in order to investigate whether or not there are differences between them in their ability to recognise isolated words in the Russian input. If there is a robust effect of metalinguistic knowledge, which is commonly attributed to linguistically sophisticated participants due to the training in linguistics they have received, it can be predicted that linguistically sophisticated participants will perform better on all measures in the present thesis.

1.4 Significance and research questions

To sum up, the points discussed in Chapter 1, the present study is significant for several reasons. Firstly, it is the study in the *ab-initio* learners' paradigm, which adds to paradigm knowledge about how words are detected in a new language pair, such as L1 English and L2 Russian. Secondly, the present study was designed specifically to investigate the effect of such cues as phonotactics, prosody, word length, and their interactions have on participants' ability to detect words in Russian. Thirdly, the study was designed to examine the effect of L1 knowledge, such as L1 phonotactic transfer and Metrical Speech Segmentation strategy (MSS) (discussed in Chapter 4) when detecting words of Russian, and whether or not *ab-initio* learners could generalise the phonotactic properties of words they heard in the input to new items. Fourthly, the present study uniquely examines whether or not there are differences between linguistically-sophisticated and linguistically-naïve participants in word detection abilities concerning the aims mentioned above.

Furthermore, previous research on *ab-initio* learners, for instance by Gullberg *et al.* (2010, 2012), Rast (2010) and Shoemaker and Rast (2013) has found an effect of an increasing amount of input on the ability to extract words from unfamiliar input (discussed in Chapter 3). Also, a study by Davis *et al.* (2009) showed that new words learnt the night before testing become consolidated in the memory. This was evidenced by the slower response to their base words as opposed to control items learnt on the same day, so that this effect can be found only after sleep. As a result, the present study examined the effects of exposure to Russian, which was limited to seven minutes on each day over four consecutive days. It additionally examined how learners develop over time by relying on cues such as phonotactics, prosody, length of words, and their interactions.

Finally, as will be seen in Chapter 3, one of the limitations of studies on *ab-initio* learners is that not a single study has controlled the extent to which participants were paying attention to the input when learners were going through the input phase. Lack of attention could have affected the results. All studies, which tested if there was any effect of true cognates, found strong evidence that cognates were recognised very well when presented in isolation (Rast 2010; Shoemaker and Rast 2013) and even in sequential contexts (Carroll 2014). Therefore, in addition to utilising a word recognition task and a forced-choice task (both of which are common in word-learning experiments designed to test word learning or word detection), the present study also uses a cognate identification task (see Chapter 5). The purpose of this task was to eliminate from the analysis the responses of those participants whose performance on the cognate identification task was low.

The present study aimed to answer the following research questions (RQs):

1. Does learners' ability to detect Russian words from the input increase over sessions?
2. Do learners rely on L1 phonotactics, or they develop sensitivity to Russian phonotactics when detecting words from the input?
3. Do learners rely on MSS (strong-weak stress pattern), or they rely on weak-strong stress pattern when detecting words from the input?
4. Do learners show preference to bisyllabic over monosyllabic words when detecting words from the input?
5. Are learners guided by an interaction of phonotactics and MSS when detecting words from the input?
6. Are learners guided by an interaction of phonotactics and word length when detecting words from the input?
7. Does sensitivity to phonotactic constraints in the detection of words from the input increase over sessions?
8. Does sensitivity to MSS (strong-weak stress pattern) in the detection of words from the input increase over sessions?
9. Does sensitivity to word length in the detection of words from the input increase over sessions?
10. Can learners generalise phonotactic properties of words heard in the input to new words?

11. Is there a difference between linguistically sophisticated participants and linguistically naïve-participants in their ability to detect words from the input with respect to every hypothesis (1-12) which are formulated above.

1.5 Overview of the thesis

The main aim of the present study is to investigate whether or not *ab-initio* learners of L2 Russian can detect words in this new language after minimal input and if the detection of words improves after four sessions of aural exposure to Russian. More specifically, it aims to investigate if these *ab-initio* learners can rely on phonological cues (phonotactics and prosody) and a distributional cue (word length) for detection of word boundaries and whether they can generalise knowledge of phonotactics to new items. For this purpose, Chapter 2 presents a detailed account of how these cues, among other phonological cues, have been shown to influence speech segmentation in L1 infants, children and adults. Additionally, as the present study also aims to examine if *ab-initio* learners of Russian are guided by native phonological constraints such as phonotactics and prosodic patterns, and whether or not they are sensitive to phonotactic constraints in Russian when detecting Russian words, Chapter 2 additionally reviews the literature on L2 speech segmentation with a focus on those studies which look at learners at stages beyond minimal input. Chapter 2 describes MSS, in Section 2.2.4, and the sonority sequencing and minimal sonority principles, in 2.2.5. Chapter 3 is specifically dedicated to the discussion of studies on first exposure, which incorporate phonological and distributional cue aspects in word segmentation after minimal input. Chapter 3 additionally aims to explain what an *ab-initio* learners' paradigm involves. Chapter 4 then summarises the phonologies of English and Russian and compares their phonemic inventories, phonetic processes, phonotactic constraints and stress patterns. Chapter 5 provides a comprehensive description of the methodology of the present study: information about participants, experimental stimuli, and experimental tasks. The hypotheses of the present study are listed in Section 5.3. Next, the results of the present study are summarised in Chapter 6. Chapter 7 is dedicated to the discussion of the results and the major findings. Chapter 8 then present the general conclusions of the present study in light of its limitations and gives suggestions for future research. The description of the materials which were used in the present study can be found in the Appendices.

Chapter 2. Studies on speech segmentation of L1 and L2 learners

2.1 Introduction

We saw in the previous chapter that speech stream is a continuum, and it does not provide discrete meaningful units (also known as '*words*'). Language, in contrast, is perceived and processed as units (Carroll 2012: 230). Therefore, language learners must convert a continuous speech stream into units and must create a representation of how these units are sequentially and systematically related (Lust 2006: 143). Nowadays we know more and more about infants' segmentation abilities, but we know very little about how L2 learners make the transition from the stage of hearing incomprehensible noise to the stage where she can hear some sequence of syllables (Carroll 2004: 236). In an attempt to fill this gap, we need to investigate the segmentation abilities of L1 learners, and adult *ab-initio* L2 learners, and of course we need to understand what happens in segmentation abilities of adults who accumulated more than minimal experiences with a language, whether L1 or L2. The purpose of this chapter is to review the existing research on the role of phonological cues and distributional cues in segmentation abilities of L1 learners (including infants, children and adults) and adult L2 learners who accumulated significant experience with their L2. This prepares a foundation for the discussion of studies on *ab-initio* L2 learners' segmentation abilities, which is in Chapter 3.

The indications are that babies come into this world with some basic speech perception capacities as it is evident in infants' unique ability to discriminate a wide range of speech contrasts, for example, they can differentiate between the sounds [ba] and [pa] (Eimas *et al.* 1971), different speaking rates (Eimas & Miller 1981), and talkers' voices (Jusczyk *et al.* 1992; Kuhl 1985) within the first few months of life, and even before they start producing language. These basic speech perception capacities equip infants with a foundation for finding out about what is possible and what is not in their native languages. Perhaps, the first manifestation of this ability (that is, indication of some familiarity with a native language) is documented in research by Mehler *et al.*

(1988) who found that French infants only 4 days old could distinguish speech samples in French from Russian, and that 2-month-old American infants could discriminate speech samples in English from Italian, each speech sample in both groups was read by the same speaker who was fully proficient in two languages. These findings could also reflect the fact that babies experience with their native language, in fact, starts some time before they are born, as it is evident from neonates' ability to prefer their mothers' voices as opposed to the voice of another female (DeCasper & Fifer 1980). This was also evident in another experiment when neonates with an average age of 55.8 hours preferred listening to the passages which their mothers read during the final six weeks of pregnancy as opposed to control passages (DeCasper & Spence 1986).

In Chapter 2, many more studies on how L1 learners undertake the segmentation of their native languages than the existing studies on experienced L2 learners' segmentation abilities are discussed. Firstly, this is because there is more research which has been conducted on the role of various cues for segmentation in L1 than on segmentation in L2. You will see from this chapter, the studies on L1 demonstrate that infants within the first year of life learn a tremendous amount about their mother tongues. In particular, because substantial research are reviewed which examined the roles of various phonological cues which potentially aid language learners in deciding where are beginnings and endings of words. The most substantial evidence suggests that these segmentation abilities in children develop sometime between six and nine months of age. Some of the cues that infants use for speech segmentation is their knowledge about familiar sound combinations and words, allophonic variation, information about distributional probabilities, prosodic patterns, phonetic and phonotactic patterns about their language, as well as an integration of some of these cues. Secondly, the chapter would not be complete without mentioning what is known about L1 segmentation strategies because there are peculiar similarities and differences between L1 learners and adult L2 learners, which are discussed next.

As for the differences, firstly, L1 learners and L2 learners differ in their *initial state*[8]. On the one hand, there are infants, and when they are born, they have not developed any language system until they have accumulated sufficient experience with their L1, so the child has to create a language system from the input. On the other hand, there are adults, and by the time when they are exposed to a new language for the first time, they have already developed at least a system of their L1 which in most cases has a profound influence on L2 acquisition through the negative language transfer. Secondly, many studies demonstrate that children can perceive language in the first year of their lives and before they can produce their first words (e.g. Jusczyk & Aslin 1995 among many others discussed in this section). Smith (1973) and Swingley and Aslin (2000) demonstrated that when children start producing their first utterances, their perception surpasses production of language which shows that children have what can be called "adult-like" representations about language despite not adult-like language productions. Additionally, infants between six and eight months old have an ability to discriminate sounds of a non-native language which are similar to their native language such as [ʈa]-[t̪a] for English speaking infants, whereas older English-speaking infants, children and adults cannot discriminate this contrast (Werker & Tees 1983). These studies demonstrate that infants' perception of language proceeds production. However, when it comes to adult L2 learners' perception-production, the timeline is not necessarily the same. Although, growing evidence suggests that accurate production of L2 sounds occurs only after L2 sounds are accurately perceived (Flege *et al.* 1995), there is some evidence, for instance from Japanese learners, who can produce the English /r/-/l/ contrast despite problems in perception of this contrast (Sheldon & Strange 1982), which shows that perception does not have to proceed production for L2 learners. Finally, most typically-developing children start sounding like native speakers by the time they become 3-year-old, and everyone invariably becomes a native speaker of at least one language provided s/he has been exposed to a language within a *Critical Period*

[8] In the present study, the term "Initial State" refers to state of being prior to experience (see Lust 2006: 31 for discussion of L1 initial state; and Vainikka and Young-Scholten (1996, 1998) and Schwartz and Eubank (1996) for discussion of L2 initial state.

(Lenneberg 1967). Whereas there is variability concerning the end state of those L2 learners who started to acquire language after puberty but in most cases, L2 post-puberty learners do not reach the native speakers' language mastery. For instance, a study by Patkowski (1980) showed that pre-puberty learners' accents received higher scores, in fact, those scores were similar to native speakers in an accent rating task than post-puberty learners' scores (but see Flege *et al.* 1995; Flege 1999)[9]. In this chapter, some studies show that post-puberty learners can acquire L2 segmentation strategies just like native speakers do (Weber 2000; Altenberg 2005b; Weber & Cutler 2006) in some cases due to the positive transfer (Altenberg 2005a), but other studies demonstrate that the segmentation task for L2 learners is difficult because they are affected by their L1 due to the negative transfer (Altenberg & Cairns 1983; Archibald 1992, 1993; Hart 1998; Weber 2000; Weber & Cutler 2006). It is discussed that type of task which L2 participants took can influence the interpretation of results.

Surprisingly, there are similarities between L1 and L2 learners. Firstly, both learners just when they encounter with a language for the first time have no words of a language, and knowing words would seem to be a prerequisite for virtually every other achievement of language learning. Secondly, these two types of learners are similar because before their first words are produced, both learners need to solve several important tasks. This task was discussed in different literature for children L1 learners and adult *ab-initio* L2 learners, but as you will see, these tasks are essentially the same for both types of learners. For example, it was discussed in Lust (2006:143) and Echols (1993) that the most important tasks children must do is (1) identify and extract words from the speech stream in order to map to a language knowledge; (2) store the phonological representation for a word, and relate it to a particular meaning; (3) then access that representation and construct a production from it. It was discussed in Klein (1986) for L2 learners that a learner

[9] Phonological aspect of language appears to be the most susceptible to Critical Period, e.g. Johnson & Newport (1989), Patkowski (1980); and there is some evidence that syntactic aspect of language appears to the least susceptible (Snow & Hoefnagel-Höhle 1978). Moreover, there is evidence that Critical Period does not exist at all, e.g. a study by Ioup *et al.* (1994).

needs to segment the stream of speech into discreet units (words) and to find a corresponding meaning to those words, which is followed by production attempts.

It should be clear from the above that research on L1 segmentation abilities, especially that on infants' segmentation abilities, can enhance our understanding about *ab-initio* L2 learners, and it can also help in constructing experimental designs which can be used to test our hypotheses.

Most of the research which is going to be discussed on L1 was conducted by a professor of psychology and cognitive sciences Peter W. Jusczyk who was a pioneering researcher working along with his colleagues to discover how and when language develops in babies. In his Infant Language Research Laboratory at The Johns Hopkins University, infants in most cases were aurally presented with various audio-recording of language. Depending on how old infants were, Jusczyk and colleagues measured the extent to which infants paid attention to recordings of language stimuli by several experimental methods. Among the most highly used methods are *High Amplitude Sucking* (HAS) and *Head Turn Preference Procedure* (HTPP). HAS is normally used with children under 4-months of age, who are offered a sterilised pacifier which measures infants average sucking amplitude. HTPP is normally used with infants who are older than 4-months old as they are required to sit on the laps of their caregivers for the duration of an experiment, while two sounds are played to the left and right of the infant. Infants longer looking times to the right or left while stimuli are played are taken as indications of their preferences.

The chapter is broken into six sections. Each section is dedicated towards a discussion of a particular phonological cue for speech *segmentation*, where the research by L1 (infants, children and adults) and L2 adult learners who accumulated more than just an initial experience with their L2 are reviewed. As the whole Chapter 3 reviews studies on *ab-initio* learners. The studies which are going to be discussed in this chapter are going to be studies on natural and artificial languages. Section 2.2.1 begins with an overview of studies which demonstrate that English infants can use their knowledge about familiar sounds and words for speech segmentation between six and seven and a half months of age, but it is difficult for them to extract monosyllabic words after brief input with Mandarin, which is likely to be due to the nature of Mandarin phonology. Section 2.2.2 discusses acoustic-phonetic cues for segmentation to familiarise the reader with the existence of such work. It also shows that infants and adults can rely on phonetic cues for speech segmentation. Although phonetic cues are not directly relevant to the design of the present study, an effort was made to eliminate the effect of these cues as confounding factors in the present study. Section 2.2.3 reviews studies on the role of the distributional cues on infants and adult learners' ability for segmentation. The role of distributional cues is important for this thesis because the length of words was a variable in the present study as participants were tested on monosyllabic and bisyllabic words. Section 2.2.4 is dedicated towards a discussion of prosodic cues for speech segmentation, where in addition to the explanation of MSS, studies on L1 infants and adult learners, and L2 learners are discussed. Furthermore, Section 2.2.5 lists the studies on the effect of phonotactic cues for speech segmentation. Also, it explains sonority sequencing principle and minimal sonority distance principle. The discussion of prosodic and phonotactic cues is particularly important because the present study examined the effect of stress and phonotactics in word detection after aural exposure. Section 2.2.6 reviews studies on the integration of more than one cue and their effect on speech segmentation in L1 infant and adult learners, as to the knowledge of the author of the present thesis, there are no studies on the combination of more than one phonological cue for speech segmentation in L2. Finally, Section 2.3 provides a summary for the points discussed in this chapter leading to the beginning of predictions of the present study.

2.1.1 Detection of words in fluent speech

It was decided to start this chapter with three studies which do not directly investigate the means by which or how infants detect words in the continuous speech stream, which is a more complex issue and are discussed in the next sections, instead these three studies were chosen because they ask a more general and straightforward question, i.e. whether infants *can* recognise monosyllabic words (i.e. the simplest form of words) in the speech stream after they heard these words before in isolation or whether infants can detect isolated words after they heard them before in continuous speech. Moreover, two studies by Tsay and Jusczyk (2003) and Newman *et al.* (2003), which are described in this section are on exposure to Mandarin language. These studies lend themselves to a comparison with another study on *ab-initio* exposure to Mandarin by Dutch native speakers (Gullberg *et al.* 2010; Gullberg *et al.* 2012) which are described along with other studies on first exposure to a foreign language in Chapter 3.

2.1.1.1 Studies on L1

Jusczyk and Aslin (1995) carried out four experiments to investigate how English-speaking infants detect sound patterns of English words in fluent speech. For their experiments, they selected four monosyllabic target words 'feet', 'cup', 'dog' and 'bike', and they constructed four passages with these words, each passage consisted of six sentences with the target words appearing in different sentential context to eliminate the effect of a sentence position.

In the first experiment, they employed 7.5-month-old infants where half of the infants listened to words 'cup' and 'dog', and another half listened to words 'feet' and 'bike'. After the familiarisation phase, infants listened to all four passages using HTPP. The results showed that infants had statistically longer listening times for the passages which contained familiarised words than the passages with unfamiliarised words. After that, Jusczyk and Aslin reversed the experiment by exposing 7.5-month-olds to the same passages with words (which they listened to during the test before) first and then tested them on recognition of those words in isolation. The results showed that infants had significantly longer listening times to those isolated words which they encountered before in the passages than those isolated words which they did not hear in the passages before testing. This suggests, that 7.5-month-olds can detect those words in the passages they were exposed to in isolation, as well as they are able to recognise isolated words after hearing them in the passages, before testing, which can be considered somehow more difficult than detecting isolated words. Besides, this ability appears to be specific to 7.5-month-olds because when the first experiment was repeated with 6-month-olds, there was no statistically significant difference in the infants' preference of listening to the passages. Jusczyk and Aslin, in another experiment, had 7.5-month-olds listen to items which differed from target words ('cup', 'dog', 'feet' and 'bike') by a one word-initial segment, that is 'tup', 'bawg', 'zeet', and 'gike' before they were tested on the original passages. The results showed that infants did not listen significantly longer to the passages containing original words. Considering everything, 7.5-month-old's ability to listen longer to the passages which contained words infants heard before or *vice verse* reflects that infants knew something very specific about the sound properties of their native words.

2.1.1.2 Studies on L2

In another experiment, Tsay and Jusczyk (2003) investigated whether this ability of 7.5-month-olds to segment the fluent speech can be found when the same English-speaking infants are exposed to an unknown language. To test this, they employed English learning 7.5-month-old infants who were exposed to monosyllabic Mandarin Chinese words using the same procedure as in Jusczyk and Aslin (1995). That is, they exposed half of the infants to the lists of words *tou* 'head' and *bei* 'cup' and the other half to other lists of words *dan* 'egg' and *tian* 'sky/day'. After familiarisation, infants listening times were measured while they were listening to the four passages each containing six sentences with these words which were embedded in different sentence positions. The results showed that the difference in listening times to passages containing familiarised words as opposed to passages containing unfamiliarised words was not statistically significant. On top of that, Tsay and Jusczyk replicated the experiment with four Chinese infants approximately 7.5 months of age. They found that the difference in listening times was significant this time, but due to very small sample size, they suggested to take these findings cautiously.

Additionally, the study by Newman *et al.* (2003) utilised the same procedure as in Tsay and Jusczyk (2003) with the only difference of adding of approximately five hours of video exposure (in a form of cartoons) to Mandarin Chinese which parents were asked to show to their English-speaking children two times per day for five days before visiting the laboratory with the aim to investigate whether extended exposure would aid the segmentation of Mandarin. The results showed that infants did not listen significantly longer to passages with familiarised words than to those without.

As a whole, it appears that 7.5-month old infants can segment monosyllabic words in the fluent speech of their native language, whether it is English or Mandarin, but as the studies by Newman *et al.* (2003), and Tsay and Jusczyk (2003) showed this segmentation appears not to be possible in a new language (Mandarin) this could be because exposure to Mandarin was not sufficient for English-speaking infants to start detecting Mandarin words, but also because of a number of reasons associated with a phonological aspect of Mandarin Chinese which are listed below.

1. Not all Mandarin words are monosyllabic. There are frequent words which contain more than one syllable, e.g. *gonggongqiche* “bus”, *zidian* “dictionary”, *pingguo* “apple” and others, where one syllable corresponds to a single morpheme (Duanmu 2000). It means that any monosyllabic word, in addition to being a word on its own, can also form part of longer words. Studies on distributional cues in artificial languages (e.g. Saffran *et al.* 1996b)[10] showed that polysyllabic words have higher transitional probabilities than monosyllabic words, and therefore are easier to segment. Additionally, a study by Gullberg *et al.* (2012) and Gullberg *et all.* (2010) showed that Dutch students detected bisyllabic but not monosyllabic words after brief exposure with Mandarin Chinese.
2. Each morpheme in Mandarin has a tone, and unlike English, stress is not relevant. Many words are homophones with each other, e.g. *shui4jiao4* “sleep” vs *shui3jiao3* “dumplings”. It can be expected that just a word form, which has more than one meaning can pose difficulty for language learners.

[10] The work by Saffran and colleagues on distributional cues is discussed in Section 2.1.3.

3. Mandarin syllable structures allows velar and alveolar nasals to appear in the coda position (e.g. *dan* “egg”) or either a glide (e.g. *kuai* “quick”) or a long vowel (e.g. *tou* “head”) Moreover, Mandarin syllable structure does not allow consonant clusters either in the onset nor in the coda position. It means that that the final edge of a syllable/morpheme contains only sonorous sounds which are not as useful markers of syllable edges as stops (see Sonority Sequencing Principle discussed in Section 2.1.5)

Therefore, it can be suggested that the reason why English-speaking infants could detect monosyllabic words ‘feet’, ‘cup’, ‘dog’ and ‘bike’ and they could not detect Mandarin words *tou*, *bei*, *dan* and *tian* because the final edge of English words is occupied by a stop, whereas it is occupied by sonorous sounds in Mandarin words. Moreover, we do not know what Mandarin input contained, e.g. the word *tou* depending on a tone can have twenty different meanings and it can form part of many polysyllabic words, e.g. *toufa* “hair”, *diantou* “to node”, so it is conceivable that hearing the same words pronounced with different tones, as well as hearing words on their own and part of other words could have impeded recognition of Mandarin words by English-speaking infants. However, as already mentioned, it is also possible that the reason why Mandarin words were not detected by English-speaking infants but English words were detected is because English-speaking infants did not receive as such a long exposure to Mandarin as they did with the English language. We know that a small sample of Mandarin-speaking infants detected Mandarin words which English-speaking infants could not. I discuss the other two studies on exposure to the Mandarin language by adults who were native speakers of Dutch, which was conducted by Gullberg *et al.* (2010, 2012) in Chapter 3. The next sections within this chapter describe studies which directly investigated how language learners segment words from the fluent speech. The following Section 2.1.2 describes how allophonic cues can be used for detection of words in the speech stream.

2.1.2 Context-sensitive allophonic cues

Allophonic cue is a source a language learner could use for speech segmentation. Allophony is a variation in the acoustic realisation of a phoneme depending on its phonological context. For example, in many varieties of English, aspiration is defined as a delay in the onset of voicing is considered to be an allophonic feature. Aspirated and unaspirated stops are allophones of voiceless stops phonemes in English. Aspiration is found on voiceless stop consonants when they are found in the beginning of a stressed syllable with the exception when it is found after [s], and it is not found in syllable-initial voiced stops (Davenport & Hannahs 2010). For example, ‘pin’ [pʰɪn] has got an aspirated voiceless stop, whereas ‘spin’ [spɪn] has got an unaspirated voiceless stop, and ‘bill’ [bɪl] has got an unaspirated voiced stop. Aspiration can be a helpful cue for the learner of English language to differentiate voiced from voiceless stops.

2.1.2.1 Studies on L1

Hohne and Jusczyk (1994) tested whether infants can attend to the properties of allophonic variation in English. They proposed that an allophonic contrast in English can be indicative of either a boundary between two words, or the absence of such a boundary. This contrast is finely represented in two English words ‘nitrate’ versus ‘night rate’ because the phonemic transcriptions of these two words are indistinguishable except for boundary markers /t/ and /r/ segments in ‘night rates’ which signal that ‘night’ and ‘rates’ are different words. By way of explanation, in ‘nitrates’ /nʌɪˈʈʰ r̥eɪt/, [ʈʰ] is aspirated, released and retroflex, [r̥] is devoiced indicating that both of these segments cannot be signalling word edges, therefore they must be found word-internally in English; whereas in ‘night rates’ /nʌɪt̚ reɪt/, [t̚] is unreleased, unaspirated, and not retroflex, and [r] is voiced indicating it is an initial syllable of the next word.

They employed 2-month-old English learning infants who were tested with a high-amplitude sucking procedure (Jusczyk 1985) on their ability to discriminate allophonic distinctions. They found that infants were able to discriminate ‘nitrates’ from ‘night rates’ which suggests that 2-

month-old infants can use acoustic distinctions provided by allophonic cues to know whether there is a word boundary or there is not.

In another experiment, Jusczyk *et al.* (1999a) investigated whether older infants could rely on the same information for speech segmentation by running out four experiments, adopting the headturn preference procedure from the study by Jusczyk and Aslin (1995). They used the same words as in Hohne and Jusczyk (1994) 'night rates' and 'nitrates', and two control words 'hamlet' and 'doctor'.

In the first experiment, they tested 9-month-old infants, half of which were exposed to 'night rates' and 'doctor' and another half were exposed to 'nitrates' and 'hamlet' in the familiarisation phase, and then they were tested on four passages each containing six sentences with these targets words which appeared in a different sentence position. They found that infants listened significantly longer to those passages which contained familiarised control words, 'doctor' or 'hamlet', but they did not listen significantly longer to neither the passage containing 'night rate' nor to the passage containing 'nitrates'.

Jusczyk and colleagues suggested these findings could be due to the very similar phonetic properties of 'night rate' and 'nitrate' which would demand a greater processing effort of a speech stream processing as opposed to processing the passages containing 'doctor' and 'hamlet'. The study by Jusczyk and Aslin (1995) showed that 7.5-month-old infants can segment monosyllabic words in the speech stream. Therefore, in the next experiment, Jusczyk and colleagues changed their stimuli, so they became monosyllabic items to eliminate the confounding effect of memory demand. This time they exposed 9-month-old infants to isolated words 'night' and 'dock', after that, infants were tested on recognition of these words in the same passages from the previous experiment which contained 'nitrates', 'night rates' and 'doctor', and a new passage with six sentences containing 'doc' in different sentence positions. The results showed that infants listened significantly longer to the passages which contained 'doc' than 'doctor', but the listening times for the passages containing ‘nitrates’ and ‘night rates’ did not differ significantly. Jusczyk and his colleagues concluded that reducing processing load during the familiarisation phase did not help infants’ ability to use allophonic cues for speech segmentation. They suggested that these results could be an effect of the distributional context (which is also discussed in the following sections in studies by Saffran (1996a) and Jusczyk *et al.* (1999b) in which ‘night’ always followed by ‘rates’ in the testing phase, and in fact overwhelmingly followed by ‘rates’ as it was found in both passages (the one with ‘nitrates’, and another one with ‘night rates’) which could have influenced infants in deciding that ‘night’ and ‘rates’ form a single unit.

Therefore, in the third experiment, they changed the passages which contained ‘night’ in such a way that a target word always followed by a new context, for example ‘night caps’, ‘night games’. They reasoned that if distributional cues are operating, then infants should listen longer to the new passages than the one containing ‘nitrates’. 9-month-old infants were tested on isolated targets ‘night’ and ‘dock’ and then were tested on new passages containing ‘night’ following a new word, and passages with ‘nitrates’, ‘dock’ and ‘doctor’. The results showed that infants listened significantly longer to the passages containing familiar words, that is they listened longer to the passages with ‘night’ and passages with ‘dock’.

However, the third experiment still did not show that 9-month-olds can use allophonic contrasts to segment speech as there was no clear evidence they could distinguish ‘nitrates’ from ‘night rates’. That is why Jusczyk and colleagues carried out the final experiment where they employed 10.5-month old infants to examine if they can use allophonic cues in distinguishing between ‘nitrates’ and ‘night rates’. They used the same design as that of the first experiment. That is infants were exposed to either isolated targets ‘night rates’ and ‘doctor’, or ‘nitrates’ and ‘hamlet’, and then heard passages containing all of these words in the testing phase. The results showed that 10.5-month-olds listened significantly longer to the passages containing familiarised words than unfamiliarised words, which suggests that an ability to use allophonic cues for speech segmentation develops between nine and ten and a half months of age.

2.1.2.2 Studies on L2

Altenberg (2005a) did a study to investigate if acoustic phonetic cues such as aspiration, a glottal stop insertion, as well as a combination of aspiration, glottal stop and a creaky voice can be used for segmentation of natural English speech into words. To test this, Altenberg (2005a) utilised English monolinguals as a control group and another group of L1 Spanish learners of L2 English at an intermediate and advanced levels of proficiency. Altenberg (2005a) used Spanish because it does not have aspirated stops, whereas in English word-initial voiceless stops are aspirated, so she predicted that L2 learners would find it difficult to use aspiration as a cue for speech segmentation. Additionally, a presence of a glottal stop can signal a word boundary in English, and a glottal stop boundary is found in Spanish in emphatic speech, so Altenberg predicted that L2 group would have fewer difficulties in using a glottal stop for segmentation of English.

All participants took part in a perception task where they heard a phrase, e.g. *chief's cool*, and participants needed to select between two options if they heard (1) *chief's cool*, or (2) *chief school*. The stimuli phrases were broken into three experimental conditions. The first is *chief's cool,* is an example where aspiration is a perceptually salient feature which provides a cue to a correct segmentation of the first option respectively. Other examples were used as stimuli in a perception task are: (1) *a nice man*, *an ice man*, where a glottal stop provides a boundary; and (2) *like old, lie cold,* where both aspiration and a glottal stop along with a creaky voice provided a boundary.

The results of this experiment showed that there was no significant difference between L2 learners at intermediate and advanced proficiency groups, and that all L2 learners used the best 'aspiration + glottal stop + creaky voice' cue for finding the correct word boundaries with the mean percentages of correct responses for L2 group 92.5%, which followed by the glottal stop 88.4%, and the least correctly participants used an aspiration cue for speech segmentation with 58.5% of correct responses.

This makes 76% on total of correct segmentation in L2 learners, while native speakers were correct 96% of the time. These results demonstrate that L2 speakers were much better at using the allophonic cue, which is present in their L1 (Spanish) than using aspirated voiceless stops for segmentation as they are absent in Spanish. Altenberg (2005a) suggested that her findings were most likely to be a result of L1 transfer, but it could have also been a reflection of the fact that a glottal stop is generally easier to acquire than aspirated stops.

2.1.3 Distributional or statistical cues

Another possible source of information which can be relied on to identify word boundaries in a language is the statistical information contained in sequences of sounds. These statistical regularities are meant to distinguish recurring sequences of sounds which are found within real words of a lexicon from more accidental sound sequences which are found between words of a lexicon. This statistical information is generally known as distributional or *transitional probabilities*[11] which are the terms honed by Saffran and her colleagues Aslin and Newport in the late nineties. There was a number of studies which investigated how statistical information influences speech segmentation (Brent & Cartwright 1996; Saffran 1996a; Saffran *et al.* 1996b; Aslin *et al.* 1998; Johnson & Jusczyk 2001).

There are *word-internal* and *word-external* transitional probabilities in a particular language. Word-internal transitional probabilities (also known as within-word probabilities) refer to a high probability of two sounds within a word to occur next to each other. Whereas word-external transitional probabilities (also known as between-word probabilities) refer to the situation when a chance of one sound to follow another within a word is low, so their occurrence next to each other must span a boundary between words. For example, in English, probability of [bi:] given [beɪ] is

[11] Distributional and transitional probabilities will be used interchangeably in this thesis.

very high, so it produces the word [ˈbeɪbi], but probability of [tu:] given [beɪ] is very low. Therefore, there must be a boundary between these two syllables, ‘bay#too’. Saffran *et al.* (1996a, 1996b) and Aslin *et al.* (1998) suggested that these difference between within-word and between-word probabilities act as cues to either word as a whole unit or boundaries between words, and this information is available and can be computed from the input by learners.

Saffran *et al.* (1996a) investigated if 8-month-old infants can segment continuous speech of a completely artificial language by means of distributional probabilities. They adopted the same strategy as in Jusczyk and Aslin (1993) by familiarising English-learning infants to a continuous speech stream for two minutes which was made of the following four trisyllabic nonsense words ‘bidaku’, ‘padoti’, ‘golabu’ and ‘bidaku’. The speech stream contained co-articulated consonant-vowel syllables with transitional probabilities as the only possible cues to possible words and boundaries between them. It contained no other phonological information which could have signalled word boundaries, such as there was completely no pauses and no stress. Transitional probabilities for within-words (for example ‘bida’) were 1.00, and for between-words (for example ‘kupa’) were 0.33.

After familiarisation phase, in the first experiment infants were tested on their ability to recognise the words from the familiarisation phase (which were the strings from the input phase) as opposed to nonwords (which were created for the purpose of the experiment, these words were made of the same syllables as in the familiarisation phase, but the strings were in a completely different order[12]). The results showed that infants listened significantly longer to nonwords then words from the input.

In the second experiment, infants were required to perform a more complex task by being tested after the familiarisation phase whether they prefer listening to the same words from the input as

[12] The transitional probabilities of none words equalled zero relative to the input as infants did not heard them in the familiarisation phase.

opposed to part-words (which were also created for the purpose of the experiment, these words were trisyllabic which was created by adding the final syllable of a word from the familiarisation phase to the first two syllables of other words from the familiarisation phase[13]). The results showed that infants listened significantly longer to part-words stimuli than the words from the input.

This dishabituation effect in both experiments shows that infants can differentiate new words, and even more difficult for recognition part-words, from the words they heard during familiarisation phase, which means that 8-month-olds have the capacities to use distributional cues for the speech segmentation of unfamiliar input after as little as two minutes of listening. Using nearly identical stimuli and testing procedure, Aslin *et al.* (1998) tested another group of 8-month-old English learning infants. Just as in the previous experiment, they found that infants preferred to listen to the relatively new unfamiliar part-words than words they encountered in after as little as 3 minutes of exposure. Additionally, Johnson and Jusczyk (2001) replicated a study by Saffran *et al.* (1996a) in which instead of synthesised speech stimuli they utilised natural speech stimuli. Johnson and Jusczyk (2001) found that infants performed similarly to that study by Saffran *et al.* (1996a) by listening longer to the novel part-words showing that they were able to discover word boundaries by relying on distributional probabilities in the natural speech.

However, it is not clear whether adults are capable of using transitional probabilities as cues to word boundaries; the next study asked this question. Saffran *et al.* (1996b) carried out similar experiments to the above one but with monolingual English adults who were undergraduate students. In the first experiment, Saffran and colleagues created another 'nonsense' language which was created of six trisyllabic words ('babupu', 'bupada', 'dutaba', 'patubi', 'pidabu' and 'tutibu') with the transitional probabilities within words from 0.31 to 1.0, and between words from

[13] For example, from two words infants were familiarised in the input 'daropi' and 'golatu', a part-word 'pigola' was created.

0.1 to 0.2. These words were put together into a text of 4536 syllables with no pauses between words and no other phonetic and phonological cues except for the distributional cues. The text was produced with an equivalent level of coarticulation by the speech synthesiser. There were three listening blocks for three minutes each.

The procedure required subjects to listen to a nonsense language with a purpose of identifying words' beginnings and endings, they were told that they would be tested on their knowledge of words from these languages after the listening is over. Participants were tested on a forced-choice task, for each trial of the task, they needed to choose between two words (one of which was a word from the nonsense language and another one was either a non-word or a part-part[14]) by deciding which of two testing stimuli sounded more like a word from a listening phase. The results showed that participants' accuracy on non-words was at 76%, and they were a bit less accurate on part-words (accuracy at 65%), performance on both conditions was statistically significant. This means that adults are just like infants can rely on word-internal and word-external cues for extracting from the speech of novel language, and they do it as quickly as only 21 minutes of exposure.

Saffran *et al.* (1996b) carried out another experiment in which they investigated an integration effect of distributional and prosodic cues on speech segmentation. They employed adult monolingual speakers of English. Saffran and colleagues adapted the nonsense language from their first experiment by changing [b] and [d] sounds with nasals in 'mupana', 'nutama', 'patumi', and 'tumimu'. Subsequently, they created three experimental conditions, in the first of which the first syllable of a word was lengthened, in the second condition the third syllables were lengthened, and in the final no-lengthening condition, only transitional cues to word boundaries

[14] Just as there were six words in the listening phase, Saffran and colleagues created six non-words and six part-words for the testing. Non-words were created in such a manner that they had sequences of syllables from a nonsense language which never followed each other in the nonsense language. Part-words were created by taking the any two syllables of a word from the language and adding them to an additional syllable, for example a part-word 'pidata' was created out of a word 'pidabu'.

were present. Additionally, they created non-word foils the first and final syllables of which were lengthened[15]. The procedure of this experiment was identical to that of the second experiment. The results showed the subjects were most accurate on the final-lengthening condition (with a mean accuracy score 80%), and the performance on the initial-lengthening and no-lengthening conditions were very similar (with a mean accuracy score 61% and 65%, respectively) with all the differences being significant. Additionally, Saffran and colleagues compared participants accuracy on syllable lengthening and their accuracy with distributional cues alone. They found that participants used final-lengthening condition produced significantly more accurately than distributional cues alone, whereas first-syllable lengthening was not more effective than distributional cues alone.

2.1.4 Prosodic cues

Another cue which is important for the speech segmentation is prosody, which has to do with elements of speech above a phoneme and are usually properties of syllables and even sequences of words. One of the questions which was asked in the thesis was whether English-speaking adults can use strong-weak stress pattern which is common in English for detection of Russian words after minimal input, that is why only those studies which are relevant to the acquisition of stress are reviewed in this chapter. Davenport and Hannahs (2010: 78) refer to stress as the prominence of a syllable which involves more muscular effort in its production; it is louder, longer and shows more pitch variation than the surrounding syllables.

In a stress language like English, syllables can be strong and weak. There is always a full vowel in a strong syllable, for example, ‘drastic’ [ˈdræs.tɪk], ‘carat’ [ˈkarət], and there is always a reduced vowel in a weak syllable, usually a schwa, for example, ‘forget’ [fəˈgɛt], or a very short form of another vowel, for example, ‘record’ (v.) [rɪˈkɔːd]) in a weak syllable. Words can be either strong word-initially or weak word-initially. A more detailed account of stress placement in

[15] Unlike the first experiment, the part-words were not used in the testing phase.

English are provided in 4.2.2. There is substantial evidence that English listeners assume that each stressed or strong syllable begins a new word in English (Cutler & Norris 1988; Cutler & Butterfield 1992; Jusczyk *et al.* 1993b; Turk *et al.* 1995; Jusczyk *et al.* 1999b). This segmentation strategy has been known as the *Metrical Segmentation Strategy* (MSS) (Cutler 1990; Cutler 1994).

Cutler and Carter (1987) examined the properties of the English vocabulary to establish lexical statistics. They examined the frequency of words in two computerised dictionaries, and they found a common pattern between them. That is, strong syllables in a word-initial position occurred on average more often than weak syllables in a word-initial position. Additionally, they examined a corpus of 190 00 words of spontaneous British conversation, and they found that out of all lexical tokens (i.e. content words which comprised of nouns, verbs and adjectives) 59.4% were monosyllabic words, and 28.2% were polysyllabic words which were stressed on the first syllable, 2.6% were polysyllables with initial secondary stress, and only 9.8% were polysyllables with weak-initial stress. These results suggest that although English vocabulary contains two times less polysyllabic words with weak first syllable than a strong first syllable, only 9.8% of these words are actually used in spontaneous conversation. These results support the effectiveness of the MSS, which predicts that each strong syllable signal beginnings of lexical words.

2.1.4.1 Studies on L1

Cutler and Norris (1988) took the findings of Cutler and Carter (1987) to directly test a model which predicts that the occurrence of a strong syllable triggers segmentation of the speech signal. They carried out an experiment where adult English-speaking listeners required to identify real words in nonsense strings. That is, for example, participants needed to spot ‘mint’ either in a second strong syllable condition ‘<u>mint</u>AYVE’, or in the first strong syllable condition ‘<u>MINT</u>esh’. Cutler and Norris (1998) predicted that ‘mint’ should be detected faster in ‘mintesh’ rather than in ‘mintayve’, due to the involvement of lexical segmentation of the second strong syllable ‘tayve’ in ‘<u>mint</u>AYVE’ if we assume strong syllables trigger segmentation; and as only ‘mint’ is strong in

‘MINTesh’, then it is the only possible segmentation, and the fastest between two stimuli. They found that responses to segmentation of mint were significantly slower in ‘mintayve’ (when both syllables were strong) with mean detection latency 1.135 ms than in ‘mintesh’ (strong, weak pattern) with a mean detection latency of 963ms. Their findings showed that what the MSS predicts is correct when adults’ participants were tested.

The subsequent studies which are discussed in this section provide evidence whether infants can rely on the model of segmentation at strong syllable.

Jusczyk *et al.* (1993b) tested whether infants showed any preference for listening to two-syllabic words of a strong-weak pattern than of a weak-strong pattern. They created lists of items which had two bisyllabic real words of English, where the first one was of a weak-strong pattern, and the second one was of a strong-weak pattern, all pairs of words were matched by the vowel which was present in a strong syllable, for example ‘comply’-‘pliant’, ‘pomade’-’neighbour’, and ‘define’-‘final’. The strong-weak lists were played on one side, and weak-strong lists were played on the other side, which was counterbalanced across subjects. They found that 9-month-old English-learning infants listened significantly longer to the strong-weak list rather than the weak-strong list.

Jusczyk and colleagues asked whether this sensitivity could show that infants were simply sensitive to the words with strong-weak pattern as this is the most frequent pattern. To test this, they exposed 6-month-olds to the same lists of words. They predicted if 6-month-olds show a preference for the strong-weak list, then it is an indication that this stress-pattern is simply more interesting to listen to than the weak-strong pattern. The results showed that 6-month-olds did not show any preference. Additionally, they tested 9-month-olds on the same lists of words which were low-pass filtered to eliminate the possibility that phonetic and phonotactic structure could have influenced infant’s decision. The results showed that 9-month-olds once again listened longer to words of a strong-weak pattern. This finding confirms that this discrimination effect is truly a reflection sensitivity to the predominant stress pattern of English words and that this sensitivity develops between six and nine months.

Another study by Turk *et al.* (1995) added to the findings of previous research by investigating the role of syllable weight on speech segmentation. In English, a heavy syllable is the one which is either closed or have a long/tense vowel, and it is always stressed, for example, the first syllable of 'bacon' is heavy, i.e. [ˈbeɪk(ə)n]. A light syllable is the one which is an open syllable which contains a short/lax vowel, for example, the first syllable of 'beckon' is light, [ˈbɛk(ə)n].

From these examples, it is evident that stressed syllable in English does not have to be heavy. Turk and colleagues scrutinised the stimuli from Jusczyk *et al.* (1993b), which showed that most words' they used as stimuli in the experiment had tense vowels in the stressed syllables. Therefore, there was a high possibility that syllable weight could have aided infants' preference for a strong-weak pattern for speech segmentation. Turk and colleagues suggested that a preference for a strong-weak pattern may have not been observed if there was a lax vowel in the stressed syllable.

Using the same procedure as in the above experiment, they tested 9-month-old infants on three experiments to establish whether there was a preference for strong-weak over weak-strong polysyllabic words, by manipulating syllable weight of strong syllables.

In the first experiment, infants were exposed to two lists of non-words of a strong-weak and weak-strong pattern, the strong syllables of which contained a tense vowel, for example [rezəl] versus [lərez]. The average looking time was statistically longer for the strong-weak list than for the weak-strong list. In the second experiment, infants were exposed to two lists of non-words of the same stress patterns, both strong syllables of which contained lax vowels, for example [rɛzəl] versus [lərɛz]. The results were the same as in the first experiment, with the longer average looking time for the strong-weak list than the weak-strong list.

In the final experiment, they used strong-weak words from experiment two, for example [rɛzəl], and weak-strong words from experiment one, for example [lərez]. They precited that if syllable weight, in fact, has an effect, non-words which have a heavy stressed syllable should be preferred to non-words which have a light stressed syllable (for example, [lərez] > [rɛzəl]). However, they found that the average looking time was not statistically significant between the two lists. Turk and colleagues concluded that the effect of MSS is undoubtedly evident, but syllable weight is not responsible for 9-month-old's preference for a strong-weak pattern over weak-strong pattern.

Moreover, another experiment by Jusczyk *et al.* (1993a), which is described in detail in Section 2.1.5 as its main focus was phonotactics, they discovered that neither 6- nor 9-month-old infants rely on prosody to differentiate English stimuli from Dutch. However, their findings could be not due to the fact the infants have no sensitivity to the prosodic characteristic of their native language, but rather because English and Dutch are similar in their prosodic patterns. To check if infants can differentiate two languages based on prosody, Jusczyk and colleagues did another experiment in which they exposed English infants to lists of English and Norwegian words because this combination of languages differed in their stress pattern. Unlike English, pitch in Norwegian words often increases towards the end of a word (Haugen & Joos 1972). They discovered that 6-month-old English infants preferred to listen to English than to Norwegian words. When the same experiment was repeated but with the same stimuli after low-pass filtering, the same results were found. Jusczyk and colleagues concluded that infants attend to the prosodic pattern of a native language before they attend to its segmental and phonotactic information, and that sensitivity to the native language phonotactic restrictions develops sometime when infants are between six and nine months of age.

These studies summarised above suggest that adults and infants can rely on MSS in segmentation words, and that sensitivity to the prosodic structure of a native language develops between six and nine months. However, these studies did not directly test how segmentation of words take place in a fluent speech. The study by Jusczyk *et al.* (1999b) directly addressed this gap. They carried out three blocks of investigations in which they studied how 7.5-year-old infants and 10.5-year-old infants can segment words from speech sequences. All experiments were inspired by Jusczyk and Aslin (1995) study which explored how 7.5-month-olds identify repeated monosyllables in the speech stream by extending it to segmentation of bisyllabic words.

The first part of their investigation consisted of six experiments in which Jusczyk and colleagues examined if 7.5-month-olds are sensitive to the strong-weak pattern for speech segmentation. In the first two experiments they utilised four words with strong-weak pattern 'kingdom' and 'hamlet', and 'doctor' and 'candle'; and four passages which contained these words in different contexts. They always used counterbalanced design, so that for example half of the infants listened to familiarised words 'kingdom' and 'hamlet' or passages containing these words, and the other half listened to unfamiliarised 'doctor' and 'candle' or passages with those words.

In the first experiment, they familiarised infants with lists of isolated strong-weak words which were later tested on four passages, either containing familiarised words or containing non-familiarised words. In the second experiment, the procedure was reversed that is infants were first exposed to passages containing either 'kingdom' and 'hamlet', or 'doctor' and 'candle', and then they were tested on recognition of those words in isolation.

They found that in the first experiment infants had statistically longer listening times listening to passages containing familiarised words, and in the second experiment, they listened statistically longer to those isolated words which they heard previously in the passages. These showed the 7.5-month-old infants attend to properties of strong-weak words by recognising them in a passage or in isolation if they were previously exposed to them in isolation or in a passage respectively.

Additionally, they proved by the third, fourth, and fifth experiments that this respondence to strong-weak pattern was not simply the reflection of infants' recognition of strong syllables, for example 'king' and 'ham', as they were neither able to detect isolated full words after familiarisation with just strong syllables, nor could they detect isolated strong syllables after familiarisation with passages containing full words, nor could they listen longer to passages containing full-words after familiarisation with isolated strong syllables. Finally, Jusczyk and colleagues showed that acoustic mismatch between isolated strong syllables (which were recorded anew for the experiment) and bisyllabic words had no impact on infants' listening times, by repeating the experiment 5 with a new version of strong syllables by excising them from full bisyllabic words. After infants were familiarised with excised strong syllables, they were tested on passages containing full words. The results were not significant. Table 2-1 below was adapted from Jusczyk *et al.* (1999b: 178) summarised the results of all the experiments in the first part.

Experiment	Familiarization stimuli	Test stimuli	Evidence of segmentation?
1	Isolated S/W words	Passages with S/W words	Yes
2	Passages with S/W words	Isolated S/W words	Yes
3	Strong syllable passages	Isolated S/W words	No
4	Passages with S/W words	Isolated strong syllables	No
5	Isolated strong syllables	Passages with S/W words	No
6	Strong syllables from S/W words	Passages with S/W words	No

Table 2-1. Summary of results of experiments 1-7 (adapted from Jusczyk *et al.* 1999b).

In the second part, Jusczyk and colleagues examined if 7.5-month-olds are sensitive to the weak-strong pattern for speech segmentation. This time, they used four words with weak-strong pattern 'guitar', 'device', and 'beret'[16], 'surprise', and again four passages which contained these words in different contexts. Just as in part one, the designs were always counterbalanced. In the seventh experiment, infants were exposed to isolated weak-strong words and then were tested on passages containing those words. The results showed that 7.5- month-olds did not listen longer to passages containing familiarised words of a weak-strong pattern. However, when they extracted strong syllables from full bisyllabic words, such as 'tar' and 'vice', and 'ray' and 'prize', and familiarised infants with these CVC words before testing, infants preferred listening to passages containing full words, isolated strong syllables of which they heard before. That was also true when Jusczyk and colleagues reversed the order of the experiment. Additionally, as it was described in Section 2.2.3

[16] American pronunciation of beret is [bəˈɹeɪ].

Saffran (1996a) and Aslin *et al.* (1998) showed that 7.5-month-olds have sensitivity to the distributional properties in their language.

Similarly, to their study, Jusczyk and colleagues tested whether infants could use distributional properties for segmentation[17]. For this, they changed the paragraphs with weak-strong words by adding a monosyllabic item after each word, so 'surprise' was followed by 'in,' 'beret' by 'on,' and 'device' by 'to'. After infants were exposed to those passages, they were tested on whether they listened longer to isolated strong syllables of weak-strong words, such as 'tar' and 'vice', or 'ray' and 'prize', but they did not. However, their listening time was statically significant when after the same familiarisation phase, they were tested on pseudowords which were created by adding a strong syllable with a monosyllabic item, for example: 'taris' or 'rayon'. This clearly demonstrates that 7.5-month-old infants cannot segment weak-strong words, but they can segment isolated strong syllables out from weak-strong words, and they even can rely on distributional properties of input to segment isolated bisyllabic pseudowords, such as 'taris' or 'rayon', when they simulate the strong-weak pattern. Table 2-2 below was adapted from Jusczyk *et al.* (1999b) summarises the results of all the experiments in the second part.

Experiment	Familiarization stimuli	Test stimuli	Evidence of segmentation?
7	Isolated W/S words	Passages with W/S words	No
8	Isolated strong syllables	Passages with W/S words	Yes
9	Passages with W/S words	Isolated strong syllables	Yes
10	Passages with W/S words and following weak syllable	Isolated strong syllables	No
11	Passages with W/S words and following weak syllable	Isolated strong syllables and following weak syllable	Yes

Table 2-2. Summary of results of experiments 7-11 (adapted from Jusczyk *et al.* 1999b).

However, in the final part of Jusczyk and colleagues' investigation, they found that unlike 7.5-month-olds, 10.5-month-olds listened significantly longer to passages containing words of the

[17] The second and final parts of this study by Jusczyk *et al.* (1999b) talks about integration of prosodic pattern and distributional properties. Nevertheless, the focus of this study is exploration of stress effect on segmentation ability, that is why it is discussed in this section, instead of moving it to Section 2.2.6 where studies on effect of multiple cues are discussed.

weak-strong pattern if they were familiarised with isolated versions of them before testing. Additionally, by this age, infants stop attending to properties of isolated strong (for example 'tar') syllables if they were previously familiarised with passages containing weak-strong words (for example 'guitar'). Finally, 10.5-month-olds appear to be able to segment weak-strong words from the passages after familiarisation, even when they are confronted with misleading, conflicting information provided by the distributional cues which 7.5-month-olds were shown before to respond to, for example, 'guitar' vs 'guitar+is'. Table 2-3 below was adapted from Jusczyk *et al.* (1999b) summarised the results of all the experiments in the third part.

Experiment	Familiarization stimuli	Test stimuli	Evidence of segmentation?
12	Isolated W/S words	Passages with W/S words	Yes
13	Isolated strong syllables	Passages with W/S words	No
14	Passages with W/S words and following weak syllable	Isolated strong syllables and following weak syllable	No
15	Isolated W/S words	Passages with W/S words and following weak syllable	Yes

Table 2-3. Summary of results of experiments 12-15 (adapted from Jusczyk *et al.* 1999b).

To conclude, it appears from these studies that response to predominant stress pattern in the native language, that is to the strong-weak pattern in English starts at seven and half months of age and plays an important role for speech segmentation. 7.5-month-olds can segment weak-strong passage at strong syllables. Additionally, a strong syllable of a WS can be a marker of a new word in fluent speech, for example, 'tar+is' because infants at seven and a half months of age are perfectly capable of relying on distributional cues for determining where the end of a word. Besides, abilities of 10.5-month-olds are akin to those of adults, as they can segment weak-strong words. Jusczyk and colleagues believe that 10.5-month-olds higher performance can be attributed to their increased sensitivity to other cues for segmentation, such as phonotactics and allophonic cues.

2.1.4.2 Studies on L2

Few studies have investigated the acquisition of stress in the second language. Perhaps, the main work which was conducted on the acquisition of L2 was by John Archibald.

Archibald (1992, 1993) investigated how L1 speakers of Polish and Hungarian acquired English stress. He used a basic research design where all subjects needed to take both production and perception tasks. For the production task, participants needed to read a list of words which Archibald created for the experiment and these words differed in their parameters settings which followed by them reading sentences out loud. In a perception task, participants listened to the audio recordings of the same words, and for each word they needed to indicate which syllable they thought has got stress. Archibald (1992) looked at the acquisition of English stress by Polish L1 learners. Polish is a stress-fixed language, with most words stressed on the penultimate syllable. English stress placement is complicated[18], but generally, researchers agree that English stress can be predicted based on a lexical class and syllable weight. Firstly, most bisyllabic nouns are stressed on the first syllable, and most bisyllabic verbs are stressed on the second syllable. Additionally, heavy syllables either with a long vowel in a nucleus (e.g. CVV) or a consonant in a coda position (e.g. CVC) attract stress. In his experiment, Archibald (1992) found that Polish learners tended to stress English nouns on initial syllables (e.g. 'hOrizon' instead of 'horIzon'), and a tendency to stress English verbs on a final syllable (e.g. 'astonIsh' instead of 'astOnish'). This suggests that learners could access the lexical class for assigning stress in English, and they generalised this strategy to words when it was not appropriate. However, he also found that L1 Polish parameters' settings were transferred into L2 English, as participants often produced English words with antepenultimate stress as if they had stress on penultimate syllable, e.g. 'cabInet' instead of 'cAbinet'.

[18] A more detailed account of English stress placement is going to be reviewed in Chapter 4.

Archibald (1993) conducted another experiment with L1 Hungarian participants. Hungarian is essentially a fixed-stressed language with the initial syllable usually being stressed. Additionally, like in English, its syllable weight is important in determining whether a syllable is stressed or not. However, unlike English, Hungarian is sensitive only to the structure of a nucleus but not rhyme. That is a syllable is stressed only if it contains a long vowel or a diphthong but not when it contains a short vowel followed by a consonant. Firstly, Archibald found that participants tended to transfer their word-initial stress pattern to English words, that is placing stress on the first syllable of a word even when it was not appropriate in English, e.g. 'Agenda' instead of 'agEnda'. Additionally, elements of L1 transfer were seen in participants' lower accuracy on words which were supposed to be stressed because the syllable was closed, e.g. 'Appendix' instead 'appEndix', and 'sInopsis' instead of 'sinOpsis'. This strategy of stressing closed syllables exists only in English but not in Hungarian, so it is not surprising participants did not make use of it. Lastly, participants were more accurate with words which were supposed to be stressed because of a long vowel, e.g. arEna, horIzon. This suggests that L2 speakers of English relied on their knowledge of L1 Hungarian that word's initial syllable and heavy nucleus receive stress.

These two studies which were done on the acquisition of L2 stress suggest that L2 learner can reset their parameters to L2 setting, but there is also evidence of transfer of L1 stress patterns (Archibald 1992, 1993)

2.1.5 Phonetic and Phonotactic cues

There are certain restrictions within languages which define the combinations and position of speech sounds in spoken words. These restrictions are called phonotactic constraints and are highly language-specific. That is when we talk about phonotactics in a given language there are combinations and position of sounds which are possible (known as *legal*), and there are some which are not possible in this language (known as *illegal*). Additionally, phonotactics are traditionally seen as *high-probability* and *low-probability*. *Phonotactic probability* is the term which was coined by Jusczyk *et al.* (1994) and Vitevitch and Luce (1999, 2004). It has been used to refer 'to the frequency with which legal phonological segments and sequences of segments occur in a given language' (Vitevitch & Luce 2004: 481). For example, in English, /ŋ/ sound (which is found in a word 'sing') is illegal in a word-initial position as it can never occur there, but it is legal in a word-final position as it is highly frequently found there (for example, 'king', 'song', 'wing'). Whereas /h/ sound is illegal in a word-final position as it is never found there, but it is legal in a word-initial position as it is rather frequently found there (for example, 'hair', hand', 'half'). Furthermore, only a subset of consonants may form syllable-initial and syllable-final clusters, and the order of consonants within clusters is severely restricted (Clements & Keyser 1983). For example, cluster /rt-/ is not a possible syllable-initial sequence in English, but it is possible in a syllable-final position in rhotic varieties of English, for example, 'sport'. Whereas in Russian, /rt/ cluster is found in both syllable-initial and syllable-final position, for example 'rtut' [rtut^j] (mercury) and 'tort' [tort] (cake).

Additionally, consonant clusters are subjects to the *Sonority Sequencing Principle* (SSP) (Selkirk 1984)[19], which defines the order of consonants in a specific syllable. Selkirk (1984) provides perhaps the most detailed scale of the SSP, which is presented in Table 2-4 below. SSP presupposes that the edge of the syllable must be occupied by the least sonorous segment, whereas the syllable nucleus must be occupied by the most sonorous segment. In fact, SSP principle applies not only to syllables of CVC type but also to onsets and rimes with more than one segment. For instance, based on a SSP, /rt/ cannot be accepted as a possible onset /rt-/* because a liquid /r/ is more sonorous than a stop /t/, but it can be a legitimate rhyme /-rt/ in rhotic varieties of English as /t/ is less sonorous than /r/ which is exactly what is needed for it to occupy syllable-final position. So, from the point of view of SSP, /rt/ in a syllable-initial position is considered as a violation of sonority; but from the point of view of language-specific phonotactic constraints, as it

[19] There was even an earlier attempt to define sonority as the 'loudness of segments' (Bloomfield 1933)

was explained in the previous paragraph, the sequence /rt-/ can occur in a syllable-initial position in some languages, e.g. in Russian, so it is legal in Russian.

	Low vowel	
	(a)	
mid vowel	(e, o)	mid vowel
high vowels	(i, u)	high vowels
liquid	(l, r)	Liquid
Nasal	(m, n)	nasal
voiced fricatives	(v, ð, z, ʒ)	voiced fricatives
voiceless fricatives	(f, θ, s, ʃ, h)	voiceless fricatives
voiced stops	(b, d, g)	voiced stops
voiceless stops	(p, t, k)	voiceless stops

Table 2-4. The sonority sequencing principle (Selkirk 1984).

Finally, there is another constraint which is important for consideration of this thesis. *The Minimal Sonority Distance* (*MSD*) (Selkirk 1984) is a language-specific constraint which specifies that the segments within a syllable must have a certain distance or be restricted from each other. The position of these segments is explained based on their relative distance on the sonority scale. To explain this point, Table 2- 5 from Broselow & Finer (1991) was adapted, which illustrates that vowels are the highest in sonority hence index 4, and obstruents are the lowers in sonority thus index 0. As it is seen from Table 2-5, every class of sounds is assigned a sonority value (index), which varies in one interval. Languages differ in combinations of these values they allow. These few examples of syllable onsets help to clarify the point: (1) Mandarin does not allow branching onsets, so MSD of a Mandarin syllable [ba] is 4. (2) Spanish does not allow MSD to fall lower than 2, which means that such sequences as /cl-/, /gl-/, and /pr-/ are legal in Spanish, but /pv/* is not. However, there are languages which allow two obstruents next to each other, e.g. /mp/ as in some African languages, or /pt-/ as in Polish, which means that these languages tolerate MSD=0. Finally, there are languages which can go even into negative values, e.g. cluster /lb/ of Russian (which was already discussed several times throughout the thesis) takes a liquid /l/ with a value 2 as a starting point and attaches it to an obstruent /b/ with a value 0, making the MSD index of the whole clusters = -2. These can be summarised that all positive MSD values are examples of sonority rises, which are common cross-linguistically; MSD values which equate to 0 are examples of sonority plateau (which are less common cross-linguistically), and negative values are examples of sonority falls (which are rare cross-linguistically).

Liquids	Nasals	Obstruents	Nasals	Liquids	Glides	Vowels
-2	-1	0	1	2	3	4

Table 2-5. Minimal sonority distance.

To sum up, some languages violate SSP, and what is legal in one language can be illegal in another language. Researchers have long been interested in how learners acquire phonotactics of their native language, and how learners of one language respond to phonotactic properties of another language. Details of English and Russian phonotactics are described in Chapter 4, as well as what predictions can be formed based on different phonotactic properties of these languages. This chapter focuses on reviewing studies which investigated the role of phonotactic constrains in infants' and adults' perception, and production in children, as well as how these constraints can be used for the segmentation of connected speech in L1 and L2.

2.1.5.1 Studies on L1

Friederici and Wessels (1993) carried out a set of experiments to find out when sensitivity to language-specific phonotactics develop and whether this knowledge can be used for the speech segmentation. To do this, they established clusters of medium frequency which satisfied word onset and word offset conditions of the Dutch language. They employed 4.5-, 6- and 9-month-old infants from monolingual Dutch families. Infants needed to listen to the lists of legal speech samples, which consisted of isolated words with legal onset and offset (for example, 'bref and 'murt'), and they needed to listen to another list of illegal speech samples (also isolated words), the illegal sequences of which were created by inserting legal onset clusters at the end of the word, and inserting legal word offset clusters at the beginning of the word (for example, '*rtum and '*febr'). Friederici and Wessels (1993) found that 9-months-olds but not 6- and 4.5-month-olds had significantly longer orientation time to the phototactically legal sequences as opposed to the illegal ones which they took as evidence of the sensitivity towards the phonotactic patterns of their native language.

In an additional experiment, Friederici and Wessels (1993) discovered that when the same words were surrounded by a word ‘mig’ from the beginning and from the end, creating the following legal onset condition sequence ‘mig bref mig’, and legal offset condition sequence ‘mig dint mig’, as opposed to illegal onset condition sequence ‘*mig ntit mig’ and illegal offset condition sequence ‘*mig feBR mig’, 9-month-olds listened longer to the legal list. Besides, this effect was present when the interstimulus interval between speech samples was reduced to 800 msec from the original 1.250 msec, and even when the speech samples were read in an infant-directed speech mode[20]. However, this effect was not found when the stimuli were low-pass filtered[21], which means that infants’ preferences were undeniably due to their sensitivity to phonotactic information. Friederici and Wessels (1993) concluded that 9-month-olds have knowledge about legal patterns of their native language, and they can use these patterns in recognition of words’ boundaries in simple sequences of speech.

In another experiment, Jusczyk *et al.* (1993a) also investigated when infants start attending to the phonetic and phonotactic properties of their native language. They created lists of low frequency abstract words in English and Dutch, which were recorded by a bilingual talker. English and Dutch were chosen because of their similar prosodic properties, so researchers could control for the influence of prosody while focusing on how phonetic and phonotactic information influences infants’ ability. They chose some words that had segments and sequences of segments which were impermissible in the other language. After infants listened to the lists of these words, Jusczyk and colleagues discovered that English 9-month-olds were able to discriminate English from Dutch stimuli, but when the same experiment was carried out with 6-month-olds, two languages could

---

[20] The characteristics of the infant directed speech mode are (1) an increase of decibels, (2) a higher pitch, (3) a lengthening of the critical items, and (4) an overall exaggerated stress pattern, (Friederici & Wessels 1993: 292)
[21] Low-pass filtering which is applied to the stimuli is often used in linguistics experiment as it preserves prosodic information but disrupts phonetic and phonotactic cues from the stimuli.

not be discriminated. Moreover, when the same stimuli were low-pass filtered, there was no statistically significant difference in infants' preference in listening to one language over the other, indicating that prosody did not affect the infant's preferences.

Jusczyk and colleagues wondered whether there was something peculiar about English words which could have attracted 9-month-olds from any language background. To test this, they redesigned their materials in such a way that they eliminated and changed all items which had phonemes unique to English or Dutch (for example, segment /Ɵ/ appears only in English thus any word with a segment /Ɵ/ was eliminated from the list with English words), consequently leaving two lists in each language to be different by phonotactics (permissible sequences in each language). After that they exposed American and Dutch 9-month-olds to that stimuli, they still found the same results, that is American infants preferred to listen to English list and Dutch infants preferred to listen to Dutch list, but the extent to which Dutch infants listened to Dutch was not as good as that of American infants listening to English. Jusczyk and colleagues concluded that it was because of the exposure of Dutch infants to English through the media (they found that they listened to English 1.25 hour a day). When they repeated the same experiment with 6-month-olds, they did not find any listening preference as with the 9-month-olds above. Additionally, when the same experiment was carried out with 9-month-olds using stimuli which were low-pass filtered, Jusczyk and colleagues found that infants did not show preference of their native language which confirms that the found effect in the previous studies was due to phonotactics. The researchers concluded that infants know a sufficient amount about phonetic and phonotactic information about their native languages to be able to distinguish their native language from another language[22].

[22] As part of Jusczyk *et al.* (1993a) experiment, they utilised another combination of languages which is unlike English and Dutch which had similar prosodic characteristics, differed in their prosodic characteristics. This language pair was English and Norwegian. It is discussed in detail in section 2.1.4.

Jusczyk *et al.* (1994) did a study with 9-month-olds who were exposed to two lists of monosyllabic non-words of a CVC structure, one of which consisted of a high-probability phonotactic pattern, and another one was made of a low-probability phonotactic pattern in English. The phonotactic probability was calculated by taking positional phoneme frequency and biphone frequency into account which were computed based on log frequency-weighted values (Kucera & Francis 1967). Infants were exposed to these two lists during the familiarisation phase, and then they were tested on the same lists. It was found that 9-month-olds listened longer to the list of a high-probability phonotactic pattern than the low-probability one. Jusczyk and colleagues thought that this result could be the reflection of the fact that high-probability items could be more interesting to listen to than the low-probability ones. To eliminate this possibility, they tested 6-month-olds on the same stimuli. The results showed no statistically significant difference in 6-month-olds preference in listening one list over the other.

The studies described above showed that infants are not only sensitive to the phonotactic patterns of their native language at 9-months of age, but they can also respond to the properties of the phonotactic probability patterns in the native language. This sensitivity appears to emerge sometime between six and nine months of age.

The next studies, which are described in this section, show how information about phonotactics can be used in finding word boundaries. Mattys and Jusczyk (2001) did a study where they directly investigated if 9-month-olds use their sensitivity to within- and between-words phonotactics for on-line word segmentation. For this, they came out with a word ‘gafe’ and a non-word ‘tove’ (both of CVC structure) because their word-initial and word-final consonants were satisfactory for researchers to create within- and between-word clusters which would proceed and end them, in the following way C.CVC.C. It was explained in Section 2.2.3, what is meant by a word-internal and word-external distributional context. In essence, within-word cluster and between-word clusters are similar to word-internal and word-external probabilities. Mattys *et al.*

(1999) defined a within-word cluster as a cluster which occurs frequently within words and infrequently between words, and a between-word cluster is a cluster which occurs frequently between-words and infrequently within words.

In the first experiment, Mattys and Jusczyk (2001) familiarised infants with one passage with phonotactics cues present, that is with the between-word cues, which had good phonotactic word boundary cue from the onset and offset to the target, for example '…brave tove trusts…'[23]; and another passage with phonotactic cues absent, in other words, without good phonotactic boundary cues, or with within-word cues, for example '… fang gaffe tine…'[24]. After the familiarisation phase, infants heard the stimuli presented on their own, that is two words they heard in the familiarisation phase 'gafe' and 'tove', as well as two control stimuli 'pod' and 'fooz' which were new. The results of this experiment showed the infants had significantly longer listening times to the stimuli which they were previously exposed to in the passage when phonotactic cues were present, which followed by the identification of targets from passages with phonotactic cues absent and two control items. Additionally, the identification of the target was the same regardless whether it was 'gafe' or 'fooz'.

In the second experiment, Mattys and Jusczyk (2001) decided to investigate whether having only word onset phonotactic cues present would be enough for segmentation of connected speech. They exposed 9-month-olds to a modified version of the passages from the experiment one. This time, in the phonotactic cues present passage, only the word onset cue was retained by having a between-word cluster, whereas the offset of the stimulus had a within word cluster. The passage with phonotactic cues absent was the same. After the familiarisation phase, infants were presented with the same four stimuli. The results of this experiment were identical to those of the first

[23] For example, the sequence '…brave tove trusts …' has between-word cues which are good for spotting 'tove' because cluster [vt] can only separate words in English, therefore it is easy to recognise 'tove' in this context.
[24] For example, the sequence '…fang gaffe tine …' has within-word cues which are bad for spotting 'gaffe' because [ŋg] ad [ft] can frequently occur between words in English, therefore making it more difficult to identify 'gafe' in this context.

experiment, that is 9-month-olds had statistically longer listening times to the items from the phonotactic cues present passage than the rest of the stimuli. This showed that having only phototactically cued onset is sufficient for spoken stimulus identification.

In the final experiment, it was tested whether having only offset phonotactic cues present is enough for the speech segmentation. The passage with phonotactic cues absent, procedure and the four lists of stimuli were identical to the two experiments described above. Whereas, the passage with phonotactic cues present was modified so that the word onset cue was absent by having a within-word cluster, and the word offset cue was present by having a between word cluster. The results of this experiment were like the above showing that 9-month-olds had statistically longer listening times to the stimuli from the phonotactic cues present passage than the rest of the stimuli. However, it is interesting to note that this effect of the offset phonotactic cues present passage is weaker than those when phonotactic cues are present from both sides of the target and onset phonotactic cues present. All in all, this study provides substantial evidence that between-word clusters can be used as a segmentation cue for extracting monosyllabic words from the connected speech by 9-month-old infants.

These experiments outlined above tell a lot about phonotactic knowledge which infants appear to internalise sometime between six and nine months of age. However, they do not tell us about the phonetic and phonotactic capacities of older learners. Subsequently, I would like to describe a production study by Messer (1967) who presented 3;7-year-old children with 25 pairs of monosyllabic words. One of each pair had a possible word of English, for example, ‘frul’ and a second pair had a pair of impossible or very infrequently used word of English, for example, ‘mrul’. In 15 out of 25 words, only initial consonants were not possible in English, for example, ‘mrul’ or ‘ʃkib’, so an example of a pair is ‘frul’-‘mrul’. In the other 10 pairs, word-initial and word-final consonants were not possible in English, for example, ‘dzrulv’ or ‘gnilb’, and an example of a pair is ‘trisk’-‘tlidk’. Children were presented with the pair of these words in a word game and were asked to judge which one sounded more like English[25].

Children’s responses were recorded manually and electronically, and they were later transcribed. Some words were discarded if they could not be evaluated. The results of the experiment showed that the phototactically-legal words were chosen more frequently as English-like than the phototactically illegal counterparts in general. Additionally, Messer (1967) found that those words phonotactics of which violated English in word onset and offset were easier to judge as non-English like than those words the onsets of which were manipulated. Finally, it was found that the impossible non-words were mispronounced more than the possible ones.

The study Brown and Hildum (1956) investigated how adults native speakers of English respond to the legal versus illegal stimuli under the conditions of noise. The stimuli were divided into three experimental conditions, (1) they were real English words, (2) phototactically legal nonsense words, and (3) phonotactically illegal nonsense words. The employed two groups of subjects, a naïve group and linguistically sophisticated group, both of which were exposed to the stimuli under the conditions of noise and we asked to transcribe what they heard. Additionally, only the sophisticated group was instructed to expect illegal items. They found that both groups of subjects were the best at identifying and transcribing the real English words which followed by their identification of phonotactically legal non-words. Brown and Hildum (1956) concluded that the knowledge of phonotactic constraints is robust and effective even when participants are told to expect illegal sound combinations.

[25] Alternatively, children were asked the following ‘which of the non-English pair (a) better described an oblong wooden block to which an experiment pointed; (b) sounded more like something he has head before; (c) sounded better to him.’ Described in (Messer 1967: 610)’

2.1.5.2 Studies on L2

Weber (2000) and Weber and Cutler (2006) conducted a study to see how phonotactic probabilities in English and German can be used by highly proficient German L1 speakers of L2 English for the segmentation of the continuous speech in English. They chose the combination of English and German languages as this pair of languages allowed for an interesting investigation into the effects of phonotactics as it is evident from the following examples. They selected two lists of English words, the first of which started with a phoneme /l/, and another list started with a phoneme /w/. The words from these lists were further embedded in the following conditions: (1) a clear boundary in both English and in German, e.g. [waʊnlɪst], as both languages do not allow onset /nl-/; (2) a clear boundary in English but not in German, e.g. [farʃlɪst], as /ʃl-/ is an illegal onset cluster in English but a legal in German; (3) a clear boundary in German but not in English, e.g. [gɔɪslɪst], as /sl-/ is an illegal onset cluster in German, but a legal in English; and (4) no boundary in either of the languages, e.g. [fuflɪst], as both languages allow /fl-/ onset.

The subjects participated in the perception word-spotting task where they were asked to spot embedded English words presented to them aurally and their reaction times, and numbers of misses were measured, for instance, a target ‘list’ as in the examples above. In addition to the experimental group of highly proficient German speakers of English, they also employed another group of native speakers of American English with no knowledge of German, which served as a control group. Weber and Cutler (2006) found that both groups of participants were affected most of all by the common boundary condition, e.g.[waʊnlɪst], which was evident in their slowest response times and the number of misses. Also, the difference between the two groups’ performance on the common boundary condition was not statistically significant. Additionally, participants were influenced by the English boundary condition, e.g. [farʃlɪst], and as before both groups performed similarly on this condition. However, only the German group was influenced by the German boundary condition, e.g. [gɔɪslɪst], that is their response times were longer and the number of misses were statistically higher than that of the English group. The results of this study suggest that highly proficient L2 learners can acquire the phonotactic constraints of an L2 and apply this knowledge for the segmentation of English words from nonsense sequences almost identically to native speakers. It was evident in German group benefiting by an English phonotactic boundary which does not exist in their native language German. However, the knowledge of L1 phonotactics (German) appears to operate even when they are listening to an L2 when it is not necessary as German group continued to be influenced by the German boundary *even* when segmenting English words which were embedded in the nonsense sequences.

Altenberg and Cairns (1983) also used two groups of subjects with English monolinguals in one group and English-German bilinguals in another group. They did a very similar study to the above. They utilised monosyllabic non-words, which were created by carefully designing word-initial consonants clusters in for the first experiment and word-final consonant clusters for the second experiment. These consonant clusters were created so they satisfied the following legality conditions; (1) legal in both English and German (e.g. *bluk* or *pelf*); (2) legal in English but illegal in German (e.g. *twoul* or *terth*); (3) legal in German but illegal in English (e.g. *pflok* or *zumpf*); and (4) illegal in both languages (e.g. *tliet* or *lepk*[26]). During the test all non-words were written instead of using aural stimuli as they wanted to be sure that subjects did not misperceive illegal sequences. Monolingual participants and a half of bilingual participants took a judgement task where they needed to rate nonwords on a scale from 1 (completely acceptable) to 5 (completely unacceptable) in terms of how acceptable they were as possible English words, and another half did the same task, but they needed to rate nonwords as possible German words. They found that bilinguals had the same responses as monolinguals in their judgment of non-words as being English-like; and that bilinguals rated non-words significantly different depending on whether those words were needed to be rated as possible English words or German words. However, when the same items were presented to participants on the screen in a lexical decision task (where

[26] These examples are not in IPA because it was not provided in the original article by Altenberg and Cairns (1983).

participants needed to press a button ‘yes’ or ‘no’ depending on whether they thought an item was an English word or not; or press buttons ‘ja’ or ‘nein’ in a German version of the test), bilinguals RTs were not the same to those of monolinguals, as monolinguals rejected faster those words which were illegal in English but legal in German. Whereas, bilingual participants were affected by the condition ‘illegal in English but legal in German’ in the same way as they were affected by illegal in German but legal in English condition. Moreover, this pattern of results for bilingual participants was the same despite the fact of whether they took an English or a German version of a test.

The results of the lexical decision task of this study are similar to that by Weber (2000) and Weber and Cutler (2006), which showed that when L2 learners took tasks where they could not use metalinguistic knowledge of their languages (i.e. when they could take time to think about their judgements or answers), their results were affected by phonotactic constraints of English and a native language German as discussed in the previous two experiments, although an activation of one of those languages was clearly inappropriate. It is interesting that these findings seem to be related to the type of a task involved to measure L2 linguistic ability to use phonotactics for recognition of possible words in a specific language, as the next study by her show.

Altenberg (2005b) did a similar experiment where she used a metalinguistic judgement task with different groups of participants. They were monolingual English speakers, and L1 Spanish learners of L2 English at the beginning, intermediate, and advanced proficiency levels. All participants took part in a metalinguistic judgment task where they needed to rate non-words as possible English words (English version was used with monolingual and L2 learners’ group), or as possible Spanish words (a version which was used only with L2 learners’ group). These non-words made three conditions: (1) consonant clusters possible in both English and Spanish (e.g. *dran*); (2) consonant clusters possible in English but not in Spanish (e.g. *spus*); and (3) consonant clusters impossible in either English or Spanish (e.g. *zban*). Altenberg (2005b) found that there was no statistically significant difference in monolingual and L2 learners groups' judgments of non-words as possible English words, that is both groups rated words like *zban* as completely unacceptable in English, and *spus* and *dran* as acceptable. However, there was a significant difference in L2 learners' performance in English and Spanish versions. That is L2 participants knew that non-words like *spus* are not possible words of Spanish but are possible in English. Additionally, there was no significant difference found between English proficiency levels.

The results of Altenberg (2005b) are identical to what was found in Altenberg and Cairns' (1983) non-word judgment tasks with highly proficient English-German bilinguals, that is participants can make judgments of what is possible in one language, and what is possible in another language, relying on what they know about phonotactic constraints in these two languages independently, that is without interference from the other. They can do it very successfully provided they are given enough time as the participants were given in the judgement tasks, but as it was shown in Weber (2000) and Weber and Cutler (2006), and in Altenberg and Cairns (1983), the phonotactic constraints of two languages can become activated in L2 learners during the different task type, lexical decision task, when one set of constraints is inappropriate in a particular language.

2.1.6 Multiple cues to word boundaries

As we saw from the previous sections above, various cues were shown to be important for the speech segmentation. That is, we know that allophonic, distributional, phonetic and phonotactic cues, as well as prosody, are reliable sources of information which infants and adults can use for in word segmentation. Although these cues as individual markers of word boundaries are undoubtfully important, Mattys *et al.* (1999) among others were the first to suggest that learning how to discover words boundaries from the connected speech is the process of knowing how to integrate these cues successfully. This section reviews those studies which investigated how infants and adults attend to properties of more than one cue for extracting words from the continuous speech.

Knowledge of how phonotactic patterns are distributed in the input could be important in isolating words from the speech input. The study by Mattys *et al.* (1999) investigated how sensitivity to phonotactics in combination with prosody can be used for the detection of word boundaries.

Mattys and colleagues selected two lists of CVC.CVC bisyllabic non-words where all C.C. sequences occurred to the same extent in connected speech, but the first list had C.C clusters of a high-probability within words but low-probability between words, for example, *nongkuth* [ˈnɔŋˑkʌθ][27]; and the second list had C.C clusters of a high-probability between words but low-probability within words, for example *nom-kuth* [ˈnɔmˑkʌθ].[28]

In the first experiment, they utilised these two lists but had all stimuli to have their first syllable stressed and the second syllable unstressed. They exposed 9-month-olds to the two lists. It was discovered that infants listened longer to the list with the high-probability within-word clusters than to the stimuli which had high-probability between-word clusters. Mattys and colleagues concluded that it was the prosodic nature of the stimuli (that is strong-weak pattern) which promoted the high within-word sequences to be perceived as a one-unit, while high between-word sequences were perceived as two-units because of the stress on the first syllable and a between word cluster which creates a conflict for a single unit perception. In fact, these findings are consistent with the studies which were previously discussed by Jusczyk *et al.* (1993b); Turk *et al.* (1995); and Jusczyk *et al.* (1999b), which showed that infants preferred listening to bisyllabic words stressed on the first syllable.

In the second experiment, Mattys and colleagues used the same two lists, but this time, they changed the stress pattern such that the second syllable was stressed. They predicted that now, having stress on the second syllable, would make the phonotactic cues of the within-word cues

[27] Based on the mother's utterances of the child-directed speech corpus, Mattys *et al.* (1999) used the following clusters: (1) high probability between, but low-within probability clusters: [ŋ·t], [f·h], [v·m], [m·h], [k·ʃ], [ŋ·b], and [m·k], [v·t], [z·n], [n·θ], [p·tʃ], [n·g].
[28] (2) High probability within, but low-between probability clusters: [ŋ·k], [f·t], [v·n], [m·θ], [k·tʃ], [ŋ·g].

conflict with the prosodic cue, thus making the perception strength less adequate. However, they predicted that the effect of the between-word cues along with the effect of prosody should be reinforced, thus promoting an easy identification of a two-unit percept. Just as with the study above, they had 9-month-olds to listen to the two lists. The results complied with their prediction, that is the lists with the between-word clusters were preferred to listen to by 9-month-olds than the lists with the within-word clusters. This makes sense, as we know that infants can use trochaic stress pattern as word's onsets markers, and high between-word phonotactics can be a cue of a new word. So for the two-unit perception, these two cues reinforced each other.

In the third experiment, Mattys and colleagues exposed 9-month-olds to the same two lists of sequences containing between and within-word clusters, but this time they inserted a 500-ms pause between the C.C syllables. They hypothesised that this boundary should act like weak-strong stress generating a preference for the list with between-word clusters. The results showed that infants had longer listening times to the list with the between-word clusters than the one with the within-word clusters. Just as in the experiment above, the results showed that strong syllables signal a word boundary, and this effect is more robust when it coincides with a between-word cluster type.

In the last experiment, Mattys and colleagues decided to test whether it is phonotactic or prosodic cues, which infants rely more when detecting word boundaries in the speech stream. To test this, they exposed 9-month-olds to the list with the within-word clusters the second syllable of which was stressed, and another list of between-word clusters which had the first syllable stressed. That was conflicting because in the first case phonotactics favour a one-unit perception, while prosody favours a two-unit perception; and in the second case phonotactics favour a two-unit perception and prosody favour a one-unit. The results showed that 9-month-olds listened significantly longer to the list with between-word clusters which were stressed on the first syllable. Mattys and colleagues concluded that when phonotactics and prosody conflict, prosodic cues have a stronger weight than phonotactic cues for the detection of word boundaries in 9-month-olds. The authors took it as another evidence about the importance of MSS for segmentation in 9-month-olds. One limitation of this study was that it did not address the issue of word segmentation from fluent speech. The study of (Mattys & Jusczyk 2001) did such an experiment where they explored the role of between- and within-word phonotactics in segmentation, which is described in Section 2.1.5. Two studies which are described next investigated how adults respond to phonotactic and prosodic cues for word segmentation.

A study by Vitevitch *et al.* (1997) was one of the first psycholinguistic studies which tested whether adult native speakers of English can use the same information which is available to infants for detection of words. In particular, they investigated whether adults can apply phonotactic information and prosodic information for speech segmentation in an on-line processing task. They conducted two experiments which were carried out with adult native English speakers too. Vitevitch and colleagues utilised nonsense syllables of CVC structure, which were of low and high phonotactic probability and were adapted from Jusczyk *et al.* (1994) experiment described in section 2.1.4. Low and high phonotactic probability was determined by calculating (1) positional segment frequency and (2) biphone frequency. For example, [kik] was a high-probability cluster, and [giΘ] was a low-probability cluster. All items, despite their probability values, were legal patterns in English. Vitevitch *et al.* (1997) had two variables, phonotactic probability pattern and stress placement. To create an experimental condition for the phonotactic probability pattern, they combined CVC syllables into a bisyllabic nonword of a CVC.CVC structure by manipulating the phonotactic probability variable resulting in four conditions, such as (1) high.high, (2) high.low, (3) low.high and (4) low.low; and in order to create a stress placement variable they subsequently stressed either the first or the second syllable of all stimuli which were generated by a previous condition.

In the first experiment, adult English speakers were tested individually or in pairs in a phonetic booth. They were presented with one of the stimulus items and were asked to rate each stimulus on a scale from 1 to 10 where 1 was a good English word, and 10 was a bad English word. Participants were given three seconds to respond, and if they did not respond for three seconds, and the null response was recorded, the trial automatically moved to the next phase. The results showed that nonsense words with the primary stress on the first syllable were judged more English-like than nonsense words with the primary stress on the second syllable. Additionally, the most English-like rated items were the stimuli which were of high.high probability and the least English-like stimuli were the ones which were the two low-probability syllables. However, there was no effect of the interaction of phonotactics and stress.

In the second experiment, they used the same stimuli as in the experiment one and the two variables: (1) phonotactic probability and (2) syllable stress were the same too. However, the procedure of this experiment was different, participants listened to spoken stimuli one by one, and after each stimulus, they were asked to repeat what they heard as quickly as possible. The reaction times from the beginning of the stimulus to the begging of the verbal response was computed. Just as with the experiment above, participants had 3 seconds to respond before the computer moved to the next trial, if no response was given, a null response was automatically calculated. The accuracy of the participants' pronunciation was measured by comparing the responses of each stimulus with their transcription. The results of the experiment showed that those stimuli which had stress on the first syllable had significantly faster reaction times than the ones which received primary stress on the second syllable. Also, words which had two high probability syllables were the fastest to repeat, whereas words which had two low probability syllables were the slowest to repeat. However, once again, there was no effect of interaction between phonotactics and stress.

The results of this study by Vitevitch *et al.* (1997) suggest that adults are like infants, in the study by Mattys and Jusczyk (2001), appear to have well-grounded intuitions about the prosodic patterns and phonotactic probabilities of their native language. This is another piece of evidence that phonotactic probability plays an important role in the processing of spoken words not only in infants but adults too. The findings that the primary stress is important is consistent with MSS (Cutler 1990; Cutler & Carter 1987; and Cutler & Norris 1988), which was discussed in detail in Section 2.1.5. Finally, as there was no significant interaction between prosody and phonotactics, Vitevitch *et al.* (1997) concluded that the effects of syllable stress and phonotactics were independent, that is in this particular study they were not facilitating each other.

Another psycholinguistic study which investigated the role of phonotactics and prosody in adults' speech segmentation was a study by McQueen (1998). However, this time, instead of English, Dutch native speakers' (*n*=52) ability to segment monosyllabic words was assessed in an on-line processing task. He chose forty monosyllabic Dutch words which appeared in the initial position of a bisyllabic sequence, for example, *pill* ('pill'), in [pil.vrem]. He embedded these words in four different contexts.

1. In the first context, a target word was stressed and it followed by another syllable which was stressed and the phonotactics between a target word and a nonsense sequence were aligned, that is it had illegal two consonants sequence after the vowel of a first syllable (StrongStrong, Aligned, as in [pi<u>l.v</u>rem]).

2. In the second context, a target word was stressed, and it followed by another stressed syllable but this time the phonotactics between a target word and a nonsense sequence were misaligned, in other words it had a legal two consonants sequence after the vowel of a first syllable (StrongStrong, Misaligned, as in [pi<u>lm</u>.rem]).

3. The third context differed from the first by being followed by an unstressed syllable which had a weak vowel schwa (StrongWeak, Aligned, as in [pi<u>l.v</u>rəm]).

4. The fourth context differed from the second by being followed by an unstressed syllable which had a weak vowel schwa (StrongWeak, Misaligned, as in [pilm.rəm]).

Additionally, further forty monosyllabic words were chosen to embed them in the final position of bisyllabic nonsense sequences, for example, *rok* (‘skirt’), in [fim.rok]. Again, to create different aligned conditions, McQueen chose different consonant sequences depending on the initial segment of the target. Each final target, just as with the initial targets (described above), appeared in four different contexts:

1. Strong.Strong, Aligned, [fim.rok];
2. StrongStrong, Misaligned, [fi.drok];
3. WeakStrong, Aligned [fəm.rok]; and
4. WeakStrong, Misaligned, [fə.drok].

Participants of this experiment were told that they would hear a list of nonsense bisyllables and that they needed to identify real words either at the beginning or at the end of those sequences by pressing a computer key as soon as they spotted a word, and then they needed to say that word aloud. Error rates and response latencies were measured.

The results of the experiment showed that participants were more accurate and faster to identify words which were aligned with phonotactic boundaries than those which were misaligned with the phonotactic boundaries for both word-initial and word-final targets. Additionally, this effect was found independently of the stress pattern of the nonsense sequences. Table 2-7 below taken from McQueen (1998: 28) shows mean percentages of error rates and mean reaction times for correct detection for all conditions. Interesting to note that participants had fewer errors and faster reaction times for those targets which appeared in the sequence final position. McQueen (1998) suggested that it could be due to the fact that when a target was found in the sequence initial position, it was the word’s offset which carried the phonotactic information about their alignment or misalignment; on the contrary when a target was found in the sequence final position, the same information was brought by the word's onset. In other words, alignment with a word's onset, for example, a word *rok* in [fim.rok] or [fəm.rok] is easier for the target word identification as it can accelerate initial access of that word, than alignment with a word's offset, for example, a word *pill* in [pil.vrem] or [pil.vrəm] which may only influence recognition of an already access word. Finally, the fact that there was no effect of stress does not go against the MSS (e.g. Cutler & Carter 1987, Cutler & Norris 1988). That is because there was not a single occasion when a target word appeared in a weak (no stress position). Therefore, McQueen (1998) concluded that phonotactic and strong-weak stress pattern should be seen as two cues which facilitate detection of words in adult native speakers of their language. These findings are consistent with the previous research.

Measure	Target position	Metrical Structure	Aligned	Misaligned
Errors	Initial	StrongStrong	32%	57%
		StrongWeak	38%	59%
	Final	StrongStrong	21%	56%
		WeakStrong	19%	63%
RT	Initial	StrongStrong	766	828
		StrongWeak	750	809
	Final	StrongStrong	535	629
		WeakStrong	499	614

Table2-6. Mean percentage missed targets (errors) and mean reaction times for correct detection (RT, in MS), measured from target-word offset (adapted from McQueen 1998:28).

2.2 Summary

The Chapter 2 reviewed studies on the role of phonological and distributional cues in L1 and L2 acquisition. As for the influence of these cues on L1 speech segmentation, we saw that infants, children and adults respond to properties of many individual cues that may facilitate identification of word boundaries.

In particular, we saw that infants as young as 2-month-old show sensitivity to acoustic distinctions provided by allophonic cues (Hohne & Jusczyk 1994), and that this sensitivity develops between nine and ten and a half months of age into an ability to use allophonic cues for segmentation of the sequential speech context (Jusczyk *et al.* 1999a). Moreover, 7.5-months old infants can identify highly specific properties of CVC words when they are presented with them again, either in isolation or in a text passage (Jusczyk & Aslin 1995). However, we do not find the same infants' abilities to identify words, when the same age English-speaking infants are presented with Chinese CVC words and then tested on them, possibly because of the phonological structure of Chinese (Tsay & Jusczyk 2003; Newman *et al.* (2003). Furthermore, there is plenty of evidence that infants and adults are sensitive to the organisation of speech sounds within their native languages. For example, studies by Friederici and Wessels (1993), Jusczyk *et al.* (1993) showed that infants develop this unique quality between 6- and 9-months of age; and at the same age infants learn to differentiate high-probability language pattern from low-probability ones (Jusczyk *et al.* 1994); and they start using these patterns to extract words from speech sequences (Mattys & Jusczyk 2001). Sensitivity to the legal as opposed to illegal sequences in a native language is also evident in children, as children are more likely to rate phonotactically legal nonsense words as English-like, and they are more likely to pronounce them correctly then those nonsense words phonotactics of which were illegal (Messer 1967). Besides, this sensitivity to native-language phonotactics is so robust that it helps adult-native speakers' decisions on what possible words of English are under conditions of noise (Brown & Hildum1956).

Furthermore, the section presented plenty of evidence that prosody can be another source of information which infants and adults can use for breaking up the speech stream. Jusczyk *et al.* (1993b) suggested that sensitivity to a prosodic pattern of a native language develops sometime between 6- and 9-months and it possibly emerges even before sensitivity to phonotactic regularities. However, at the same time, we know that sensitivity to the MSS (Cutler 1990; Cutler 1994) is present in 9-month-olds but not in 6-month-old infants, and that this sensitivity is independent of a syllable weight effect (Jusczyk *et al.* 1993b; Turk *et al.* (1995). Whereas, the study by Jusczyk *et al.* (1999b) showed that 7.5-month-olds can use strong-weak stress pattern for speech segmentation, and infants of the same age can use strong syllable of bisyllabic words as markers of potential words/nonwords when this strong syllable is followed by a weak syllable, for example: ‘tar+is’. Finally, the ability to segment words of a weak-strong pattern start to develop sometime between 7.5- and 10.5-months of age. This ability resembles that of adults, who are just like infants, were shown to be influenced by MSS (Cutler & Norris 1988). Another source of information infants and adults can use for speech segmentation relates to distributional probabilities. It was shown that 8-month-olds infants can rely on distributional properties to establish beginnings and endings of words in an artificial language which was synthesised by a computer (Aslin *et al.* 1998; Saffran *et al.* 1996a) or the same artificial language but read by a person. The same pattern of results was obtained in a replication design with adult participants (Saffran *et al.* 1996b). Last but not least, several studies were presented, which showed that infants and adults can exploit more than once cue for finding word boundaries. For example, a study by Mattys *et al.* (1999) showed that when phonotactics and prosody come to conflict, prosodic cues are preferred to phonotactics for the detection of word boundaries in 9-month-olds. A study by McQueen (1998) showed that adults detect nonsense words easier when they are embedded in a phonotactic condition which signals a word-boundary, and these nonsense words were detected better when they were aligned from an onset, rather than from the offset; despite these findings McQueen (1998) did not find a facilitating effect of stress but it still does not contradict MSS. Vitevitch *et al.* (1997) found that strong-weak stress pattern and high phonotactic probabilities influencing adult English speakers when they were either asked to provide a judgment about nativeness-like of nonsense words or repeat these words.

As for the adult L2 learners’ segmentation abilities, the Chapter has shown that adults can positively transfer their knowledge of allophonic cues segmentation strategies into segmentation of L2 English. We saw it in a study by Altenberg (2005a) with a presence of a glottal stop; however, if an allophonic cue was specific to English (e.g. presence of aspirated stops), even advanced learners experienced problems with applying this cue for speech segmentation. Additionally, with respect to acquisition of L2 stress, Archibald (1992, 1993) showed that adult L2 learners of English are greatly affected by the L1 various stress placement strategies, but he also found evidence that these learners showed sensitives to the lexical classes while stressing English words, that is nouns were stressed on the first syllables and verbs were stressed on the final syllables. Finally, the most widely studied cue in L2 speech segmentation was phonotactics. In general, studies showed that L2 learners of English can acquire English phonotactic cues and use them as effectively as English native speakers would do for segmentation of English. We saw such evidence especially with off-line tasks, e.g. judgments tasks in Altenberg and Cairns (1983) and Altenberg (2005b); and even with on-line tasks, e.g. timed word-spotting task in Weber (2000) and Weber and Cutler (2006) and a timed lexical decision task in Altenberg and Cairns (1983). However, we also saw that knowledge of L1 phonotactics appears to operate when participants were listening to an L2 and when it was not necessary. Interestingly, we observed L1 transfer only when participants took on-line psycholinguistic tasks (Altenberg & Cairns 1983; Weber 2000; Weber & Cutler 2006). Therefore, it is evident that a native-like performance exists in situations when participants took an off-line task, i.e. they had an opportunity to take time to think and to use their metalinguistic knowledge or explicit knowledge (R. Ellis 2009) to make informed judgements about language. In contrast, we observed an L1 transfer, even in advanced language learners, when they took online tasks which measured how participants processed language in time-constrained situations so that participants could not access their metalinguistic abilities or explicit knowledge. Instead, they relied on their unconscious or implicit knowledge (R. Ellis 2009) when giving a response.

To sum up, it is clear that there are multiple cues L1 and L2 learners can use for speech segmentation, and that L2 learners in many cases are biased by their L1 segmentation strategies, as it was shown above, in some cases this bias can be explained by the type of task involved. The present study is going to use psycholinguistic tasks to investigate how L1 English knowledge may influence an ability to detect words of Russian by looking at phonotactic cues, prosodic cues, and word-length cues. However, before describing the methodology of the present study, the first exposure paradigm and studies which were carried out within this paradigm are reviewed in the next chapter. It is followed by a description of phonologies of both English and Russian in Chapter 4.

Chapter 3. Studies on *ab-initio* learners

3.1 Introduction

Models of natural L2 acquisition do not take into consideration the developmental aspect of word detection when learners are confronted with foreign language input for the first time which is likely due to problems involved with controlling natural language input (see Carroll 1999, 2001, Krashen 1978, VanPatten 2000 for discussion of input in SLA). Instead, most theories of L2 learning mechanisms are based on stages during which L1 knowledge has been acquired. For example, we saw from Chapter 2 that adults are influenced by their L1 in the later stages of L2 acquisition.

As a matter of fact, researchers acknowledged more than two decades ago that too little attention has been paid to the very beginnings of the acquisition process (Perdue 1996: 138). Vainikka & Young-Scholten (1998: 31) proposed to collect data from learners at the earliest stages of acquisition in order to make claims about the L2 initial state; and researchers such as Schwartz & Eubank (1996), Pienemann (1999, 2007), Vainikka & Young-Scholten (1994, 1996, 1998), Carroll (1999, 2001) produced the earliest work where they stated their proposals on what '*initial state*' in the L2 acquisition might look like. However, this work is mainly concerned with the development of the morphosyntactic aspect of language. Therefore, it is not discussed in detail because the focus of this thesis is on how adult English L2 learners start detecting word forms from the continuous speech stream of an unknown language (Russian).

In recent years, there has been a growing line of research with the aim of investigating what the learner brings to the L2 at the initial stage of its acquisition, e.g. Rast (2008, 2010); Gullberg *et al.* (2010, 2012); Carroll (2012, 2014); Rast & Shoemaker (2013)[29]. Research which has been

[29] Only studies which incorporated phonological aspects of adult first exposure study and which focused on perception are included here, but note this list is certainly not exhaustive of all first exposure studies, see Park (2011), Carroll & Widjaja (2013), Han & Liu (2013), Ristin-Kaufmann & Gullberg (2014), Carroll & Windsor (2015) and others.

conducted within this area is known as the *ab-initio* learners' paradigm, or the *first exposure* paradigm, or the *minimal exposure* paradigm[30].

The earliest form of how a learner may make use of linguistic input and gradually approximate to the L2 was discussed in Klein's the learner's problem of analysis (1986). Specifically, he discussed that when a learner is confronted with an unknown language, s/he needs to segment the stream of speech into discrete units (words) and to find a corresponding meaning to those words. An establishment of meaning firstly goes, perhaps, with a general understanding of a meaning of an utterance, which is followed by an understating that there are separate words, each of which has meaning through the means of numerous hypotheses testing. This step is followed by the learners' synthesis problem, that is production attempts which go beyond one-word stage, which nowadays researchers would call generalisation beyond exemplars in the input to novel items and the formation (this was discussed in Gullberg *et al.* 2012 and Han & Liu 2013). The present study is not concerned with either identification of meaning or production, but it aims to investigate whether L1 English *ab-initio* learners can make use of phonological cues in L2 Russian to detect words in this new language after four sessions of aural exposure with it. The purpose of this chapter is to provide a summary of this paradigm and studies which were carried out within it. There is a comparative summary of the experiments in the final section, which follows with some predictions of the present study.

3.1.1 *Ab-initio* learners' paradigm

The present study is a study which was conducted within the *ab-initio* paradigm. It refers to a research agenda which examined what can be learned about a novel L2, the exposure to which was limited and highly controlled from an absolute onset with its encounter. We saw from Chapter 2 that adult post-puberty learners at different degrees of proficiency are biased by their L1 in the

[30] *Ab-initio* paradigm, first exposure paradigm, and minimal exposure paradigm mean the same thing in this paper, as well as *ab-initio* learners, first exposure learner, and minimal exposure learner. However, for consistency, I will be predominantly using the term *ab-initio.*

acquisition of L2 phonology. Studies on *ab-initio* learners have the potential to show whether L1 transfer operates in the precise beginning of the initial stage of L2 acquisition.

A limited number of studies have been carried out within the *ab-initio* paradigm. These studies, in general, have shown that learners show sensitivities to L2 structures for the most part, but they also showed that L1 transfer operates from the first stages of development. The main research issues which have been raised within this paradigm can be summarised in the following points, taken from Rast (2008: 29):

(1) Finding out about learners' pre-existing linguistic knowledge, such as how L1 and other L2s affect an ability to process an unfamiliar L2 (e.g. Rast 2008, 2010; Gullberg *et al.* 2010, 2012; Carroll 2012, 2014; Rast & Shoemaker 2013).

(2) Finding out about what role *implicit learning* plays in L2 learners' ability to process an unfamiliar L2 (Yang & Givon 1997; Saffran *et al.* 1996b; Aslin *et al.* 1998; Gullberg *et al.* 2010, 2012). R. Ellis (2009: 3) refers to implicit learning as a type of learning when learners are not aware that learning has taken place because it proceeds without making demands on central attentional resources. Implicit learning usually excludes any kind of instruction or metalinguistic explanations.

(3) Finding out about which role *explicit learning* plays in an L2 learners' ability to process an unfamiliar L2 (e.g. De Graaff 1997; DeKeyser 1997; Carroll 2012, 2014[31]). Ellis (2009: 3) refers to explicit learning as a type of learning when learners are aware that they have learned something because it involves memorising a series of declarative representations by putting demands on

[31] Although Carroll (2012, 2014) did not categorise her studies under explicit learning investigation, she first trained participants to remember names, and then tested them on recognition of these names in sequential context. She made it clear that during the testing phase participants were storing names in episodic memory. Episodic memory is part of explicit memory (also known as declarative memory), so it conceivable the study is likely more fitting under the explicit learning paradigm.

working memory[32]. Explicit learning usually presupposes some kind of instruction or metalinguistic explanations.

(4) Finding out about how linguistic input influences learners' processing of an unfamiliar L2 during specific language activities, for instance: perception, comprehension, grammatical analysis and production (for example Rast in her 2008 and 2010 studies managed to combine various tasks within single studies).

(5) Finding out about how much input is required and which properties of the input, L2 learners find salient (e.g. Rast 2008, 2010; Gullberg *et al.* 2010, 2012; Carroll 2012, 2014; Rast & Shoemaker 2013).

(6) Finding out about cross-linguistic reliability of the findings, from study to study, by examining different natural languages pairs. In particular, by looking at L1 and L2 language pairs which differ with respect to markedness (as defined in Section 1.2). Tables 3-1 below summarises which language pairs were studied concerning phonological markedness within *ab-initio* paradigm (e.g. in a study by Gullberg *et al.* (2010, 2012), L1 Dutch is more marked than L2 Mandarin Chinese because its syllable structure allows complex phonotactic clusters, but Mandarin does not. The same idea was applied when categorising studies in the table below.

[32] It is not entirely clear whether studies by Rast (2008, 2010) and Shoemaker and Rast (2013) looked at explicit learning. That is why I avoided putting it under explicit learning category. The studies will be discussed in detail in Section 3.1.3. One could suggest that these studies have characteristics of explicit learning because learners were instructed in Polish using a communicative approach, and such an approach does not quite resemble what learners hear in the wild. However, it could also be argued that despite the communicative approach, the Polish input excluded metalinguistic explanations of grammar and pronunciation and that is why it fits better under the implicit learning category. Regardless of the type of learning, Rast (2008) acknowledged that participants were likely to use their explicit knowledge due to the nature of tasks she utilised in all of her studies.

Study	L1 or source language(s)	L2 or target language(s)
Unfamiliar L2 is less marked with respect to the syllable structure		
Gullberg *et al.* (2010, 2012)	Dutch	Mandarin
Han and Liu (2013)	American English	Mandarin
Unfamiliar L2 is similarly marked with respect to the syllable structure		
Carroll (2012, 2014)	English	German
Han and Liu (2013)	Japanese	Mandarin
Rast (2008, 2010)	L1 French with intermediate & advanced knowledge of English, other L3 (Russian)	Polish
Unfamiliar L2 is more marked with respect to the syllable structure		
Rast & Shoemaker (2013)	French, intermediate and advanced knowledge of English, other L3 (Romance)	Polish

Table 3-7. Summaries of first exposure studies by markedness of source and target languages.

In addition to Rast's (2008) classification of first exposure study, and in addition to categorising first exposure studies by L1-L2 differences with respect to markedness, Carroll (2014: 108) classified studies which could potentially fit under the definition of a first exposure study into the following categories, such as (1) natural languages presented in laboratory settings (e.g. Gullberg *et al.* 2010, 2012; Carroll 2012, 2014; Han & Liu 2013); (2) natural languages presented in tutored conditions settings (e.g. Rast 2008, 2010; Rast & Shoemaker 2013); and (3) first exposure studies to unnatural/artificial languages presented in the laboratory (e.g. Saffran *et al.* 1996b, De Graaff 1997; Aslin *et al.* 1998; Folia *et al.* 2010; and Chambers *et al.* 2003). Some of these studies on exposure to artificial languages were described in Section 2.2 as they were relevant to the discussion of phonological cues and distributional cues for detection of word boundaries. Although, these psycholinguistic studies are highly beneficial for studying the cues presented to learners, using natural languages goes along the lines of the 'ecological validity' of *ab-initio* learners' studies (see Carroll (2014: 114) for discussion). Therefore, to act in accordance with the ecological validity, the present study, as an *ab-initio* study, is going to comply with the following three criteria:

(1) It will involve *ab-initio* learners who are genuine beginners with no experience at all of an L2 at the moment of the first encounter with it.
(2) The target language of the present study is going to be a natural language (Russian).
(3) It is going to focus on implicit learning through aural exposure to create similar conditions to what infants experience when hearing the speech stream (Gullberg *et al.* 2010, 2012).

In the next two sections, I will describe in detail a few studies which satisfied the criteria, and which also match the topic of the present study (i.e. this study investigates the effects of phonological cues). As you will see next, there are only a handful of such studies, but they represent a good example of the robustness of cross-linguistic findings from the point of different L1 and novel L2 pairs such as L1 Dutch-L2 Mandarin, L1 English-L2 German, and L1 French-L2 Polish. Additionally, these are of particular interest to the present study from the point of the tasks employed and type of input provided. I will provide a comparative summary in the final section of this chapter after scrutinising these studies, which is followed by the formulation of some predictions of the present study.

3.1.2 Studies of natural languages in a laboratory

There are only two sets of studies which investigated *ab-initio* exposure to natural languages which were presented in the laboratory and recorded by a native individual of that language. These are experiments by Gullberg *et al.* (2010, 2012) and experiments by Carroll (2012; 2014) which are described in this section.

Gullberg *et al.* (2012) employed Dutch L1 speakers who were exposed to audio-visual input in the form of the weather report in Mandarin for a maximum of 14 minutes of cumulative input as they were looking for the effect of the following variables: (1) *amount of exposure* (7 vs 14 minutes); (2) word *frequency* (occurred in the input 2 times vs 8 times); (3) *word length* (monosyllables vs bisyllables), and (4) *gesture* (highlighted vs non-highlighted) on the extraction and generalisation of Mandarin words as well as the mapping of meanings. They tested two groups of the Dutch student population with no prior experience of Mandarin or another related language on a word recognition task and sound-to-picture matching task. The first group (*n*=21) watched the video once (seven minutes of exposure), and the second group (*n*=20) watched the video twice (14 minutes of exposure).

For this, they recorded a seven-minute weather video-report in Mandarin Chinese, which was highly controlled. 24 target words were created with respect to the variables described above and were located at sentence-initial, sentence-medial, and sentence-final positions. After participants watched the movie, they were tested on a word-recognition task which consisted of target words and filler items (*n*=72) which were real Mandarin words taken from a dictionary, but participants did not encounter them during the input. All filler items were of the same syllable structure as targets. An experimental software was used to deliver a word recognition task. Participants heard experimental items one by one and needed to press a left button for *no*, right button for *yes*.

They found the amount of exposure when interacting with experimental items only slightly positively correlated with an improvement in performance, that is accuracy of a single exposure group was 55%, and an accuracy of a double exposure group was slightly higher (at 60%), with this difference being only marginally significant (p=0.05) which they took as evidence that amount of exposure alone (7 vs 14 minutes) is a not a sufficient cue for the detection of words. However, they found significant effects of syllable length and frequency variables. In particular, bisyllabic words were recognised better than monosyllabic words, and words which occurred eight times were recognised better than words which occurred two times even when the performance of a single exposure group was tested (seven minutes). Gullberg *et al.* (2012) concluded that Dutch native speakers could rely on the number of syllables (i.e. bisyllabic words) and frequency (i.e.

words occurring eight times) to identify Mandarin words in isolation at above chance level after hearing these words in sequential context after as little as seven minutes.

In another sound-to-meaning task (the details of which are not described here as it is not entirely relevant to the present study), Gullberg *et al.* (2012) found that their participants were able to match the sound structures of words they identified in the input to the referent from the input. Just as with the word recognition study, they found the effect of syllable length and frequency but also the effect of gestural highlighting, which significantly interacted with the other two predictors. In other words, high-frequency bisyllabic words which were gesturally highlighted had the highest success rate.

Gullberg *et al.* (2010) conducted another study where they investigated if adults could detect syllable structure violations of Mandarin Chinese. They used the same design as in Gullberg *et al.* (2012), that is participants firstly watched a video recording which was followed by a word recognition task in which participants needed to determine if the sounds they heard were *real Chinese.* In addition to 7 vs 14 minutes of exposure group, a control group with no exposure to Mandarin at all was utilised. The experimental stimuli were selected similarly to Gullberg *et al.* (2012), that is they were all real Mandarin monosyllabic words with the first half of them presented to participants during the video and another half was new. Additionally, there were two sets of fillers: (1) monosyllabic words phonotactics of which were violated (e.g. *gam*), and (2) monosyllabic words which comprised German-sounding clusters which were illegal phonotactics in Mandarin, word-initially (e.g. *spra*, *sna*) and word-finally (e.g., *alst*, *ans*).

The results showed that all participants were able to reject experimental stimuli, which were foils comprising German-sounding ones, even including the controls. Although the performance on monosyllabic words, phonotactics of which were violated, was 50 per cent for the group with no exposure at all to Mandarin, participants became less convinced that these words were Chinese the more input they received. Gullberg *et al.* (2010) suggested that their participants were developing sensitivity to the phonotactic structure of Mandarin in response to input, but not transferring from their L1 as Dutch indeed allows CVC syllables, because participants could identify illegal consonant sequences as not Mandarin. Finally, yet importantly, they found that participants could generalise to new items they did not encounter within the input as possible Chinese. Gullberg *et al.* (2010) concluded that Dutch native speakers could detect monosyllabic words of Mandarin they encountered before in the speech stream and to generalise phonotactic properties of Mandarin to the novel examples after as little as seven minutes of exposure.

Carroll with various colleagues since 2009 carried out a number of studies where targets in the form of German names (some of which were *cognates* with English and others were non-cognates) were presented as training trials in a laboratory setting to assess how rapidly English Anglophones with no previous exposure to German can segment these words and map them to a referent provided by a picture. Carroll (1992: 93) defines cognates as words which, when paired maybe but do not need to be semantically related, but there must be some formal resemblance between them.

I will describe in detail the most recent experiment by Carroll (2014). 50 students from the University of Calgary were divided into *beginners* in a German group with up to two semesters studying German and a *first exposure* group without knowledge of German . In a laboratory setting, participants were instructed that they would see twenty line-drawings of people (each individual was presented by two drawings to allow participants to create an abstract representation of a person) and that they would simultaneously hear twenty sentences in German. Twenty declarative sentences comprised four different structures at the end of which a target name was presented. The first task was to learn the names of the people, whose pictures they saw. This task was followed by the second task, which consisted of questions which tested if participants could detect the names they had learned before which were embedded in different phonological frames and map them to referents on the pictures.

A word list with names was used in both tasks and was created so that it consisted of German cognate and non-cognate first names and last names comprising 4-7 syllables (e.g. *Jana Langbein* or *Gisa Grunow*). Half of the last names were compounds which were semantically transparent, so they were created to pick out a referential detail from the picture. For example, the woman shown in a picture who was holding a watering can was called *Dagmar Baumgartner*, literally 'Dagmar {{tree} {gardener}},' etc. (Carroll, 2014:121).

The first task involved training trials on the list of declarative sentences which contained names form the word list in the sentence-final position, with four different sentence structures, for instance: *Hier ist Dagmar Baumgartner* 'Here is Dagmar Baumgartner'. The training trials were followed by a test of twenty questions also of four different structures, but with each structure representing a choice between two names where one was a target, and another was a foil divided by the marker *order* ('or'), for instance: *Ist hier Dagmar Baumgartner oder Trüdel Dieterich?* 'Is here Dagmar Baumgartner or Trüdel Dieterich?'.

The declarative sentences and questions (i.e. input) were recorded in such a manner that the effect of various cues could have been either controlled for or investigated. In particular, it was controlled that none of the words were focally accented, but the following variables were manipulated: for the cognates there were such variables as: (1) number of syllables (1, 2, 3 and 4); (2) number of prosodic feet in a target (one vs two); whereas for compounds there were the same variables but with different levels such as (1) number of syllables (4, 5 and 6); (2) number of feet in a target (two, three and four); (3) target word position (word-medial vs word-final); (4) syntactic frame (*istdas, istheir, sehensierhier, stehtda*) were the same; and (5) semantic transparency variable only for compound names (transparent vs opaque).

Based on the participants' success rate on training trials, they were repeated up to a maximum of nine times. The test ended when participants correctly mapped all names to pictures in the test or when participants were not able to do it after ten times. Experimental software was utilised, and an error was allocated after non-response for 2500ms. The scores were calculated automatically and were indicative if subjects could move to the testing phase to double-check that their success in the first phase was not accidental. The same questions were used in the testing (maintaining the same order of target name and foil), but a different picture of the persons' names was used. After two weeks, participants participated in the retest to measure the retention.

Carroll predicted that while the first exposure learners may segment phonetic tokens and map these to referents they would not be able to compute a morphosyntactic analysis of the compound names '*Dagmar Baumgartner*, literally Dagmar {{tree} {gardener}}' type because they have no L2 lexical entries and hence no linguistic basis yet for computing the internal structure of compound word. However, she predicted that the beginners' group would be in a position to compute a transparent semantic representation because they had some knowledge of German vocabulary, unlike first exposure group. To sum up, she predicted that beginners' group would perform better on the semantically transparent names than on non-transparent items, and better than first exposure group. Additionally, she anticipated that both groups should detect names from the input despite their length and to map them to the referents on the pictures, but she expected the beginners' group to be more accurate on this.

The results showed that participants were able to segment words of up to seven syllables and to map them to referents even on the first few items of training trials, but the beginners group required less training trials to do so on both cognate and non-cognate names, and was at an advantage over the first exposure group on compound names only, while cognates were recognized equally well. After the retest of two weeks, the beginners group was still at an advantage over the first exposure group, but identical performance on cognates disappeared, as first exposure outperformed the beginners' group. Additionally, Carroll (2014) did not find any effect of the position of the stimulus: words were segmented equally well in both sentence medial and final positions, nor any effect of a syllable structure, but she found an effect of the foot structure which she suggested to treat with caution. Moreover, she did not find any effect of the beginners' group performing better on phonologically transparent compound names than the first exposure group as both groups performed the same.

Finally, taking everything into account, these studies (Gullberg *et al.* 2010, 2012; Carroll 2014) showed that individuals can segment words from the first stages of exposure to L2 with a high degree of reliability. There will be a comparative summary of these studies in Section 3.2. after studies of natural languages in tutored conditions are discussed in the next section. Despite plenty of positive aspects of the studies discussed in this section, there are several important limitations. Firstly, the input was recorded by a single speaker in each study which does not correspond exactly to what happens when one is exposed to a language in the wild (Carroll 2001: 137). Secondly, I believe one could argue whether using training trials as in Carroll's studies on cognates goes against what happens in language acquisition the wild. It was discussed in Chapter 1 that words are not presented to infants in isolation (Woodward & Aslin 1990). Furthermore, none of these studies looked at the effect of segmentation over multiple time points, and it was not checked whether participants paid attention to the input during the exposure phase. The present study addressed the last two limitations as you will see from Chapter 5, which describes the methodology.

3.1.3 Studies of natural languages in tutored conditions settings

Rast (1998, 1999, 2008, 2010), Rast & Dommergues (2003) and Shoemaker & Rast (2013) carried out several studies where they used the language teaching paradigm using a communication-based method that excluded all use of metalanguage as well as an explicit explanation of grammar and pronunciation to expose participants to Polish. According to Shoemaker & Rast (2013) such an approach benefits from full control of the linguistic input and input treatments. This line of first exposure studies in tutored conditions settings was started by Rast (1998, 1999) herself in a pilot study where she examined the first stages of acquisition of native French 'learners' of Polish who were enrolled in the French L2 training course to become French foreign language instructors and were asked to fulfil the course requirement to study the unknown language to observe their own acquisition process. 37.5 hours of Polish input (in the form of 15 class periods) were recorded. The course was taught by a native Polish instructor. I will not elaborate on further details of these studies, as their primary aim was to find which test at the early stage could tell about *ab-initio* learners input processing. Rast mentioned that those early studies provided crucial methodological information for all her subsequent investigations, which are described next. However, these studies do not represent an exhaustive summary of Rast and her colleagues' work, as I will describe those parts of their studies which involve tasks which are relevant to finding out about L2 phonological processing. Last but not least, it needs to be mentioned that most of Rast's studies suffer from methodological limitations, that is to say from the point of participants employed, all spoke English as L2 at intermediate and advanced levels and some participants knew other languages including Slavic, which Rast and colleagues did not really account for, hence this gave rise to too many uncontrolled variables in her studies. They nevertheless raised some important points, such as types of tasks and interaction of variables.

The first study by Rast (2010) involved two groups of participants: *learners* after four and eight hours of Polish instructions, and *first exposure learners* – native French speakers who had no previous knowledge of Polish and the only input they received was that during the language task. First exposure group (*n*=34) participated in the word translation task where they were asked to read or listen to 119 unrelated Polish words and translate them as best as they could into French. She found a strong L1 influence on the translation of lexical items, first of all, based on the degree of phonetic and orthographic similarity, for example, words like *informatyke* 'computer science' (and '*informatique*' in French) were recognized well, whereas words like *rowniez* 'also' ('*aussi*' in French) were poorly recognised. As well as phonetic similarity from other L2, for example, *moi* 'my' (*'moi'* in Russian) was translated corrected by those with knowledge of Russian. However, she also found that some words, e.g. *mowi* 'he/she speaks' were incorrectly translated as 'movie', suggesting that orthographic similarity alone does not account for correct performance. She concluded that phonetic and orthographic similarity between L1 and L2 of individual words alone (without context) is not essential for the participants' ability to translate individual words (Rast 2010).

Other studies by Rast & Dommergues (2003) and Rast (2008) split participants into three time-intervals: first exposure group (*n*=8, with zero hours of instruction), and learners after four and eight hours of instruction. Participants at all time intervals were tested on a sentence repetition task where they heard 20 sentences (3-12 words long)[33] which were recorded in Polish and they needed to repeat those sentences as best as they could. They put under investigation the effects of (1) *hours of exposure /instruction* (0 vs 4 hours vs 8 hours), (2) *word length* (0-1 syllables vs 2 syllables vs 3-6 syllables), (3) *word stress*[34] (stressed words vs unstressed words), (4) *phonemic distance* (if a word contained a segment or a cluster which does not exist in French it was considered as phonemically "distant" (e.g. *nauczyciel* 'teacher') vs phonemically "close", which were all other words), (5) *transparency* (opaque vs fairly transparent vs very transparent), with transparency defined as judgments of French monolingual speakers with zero exposure to Polish who were asked to listen to Polish words and translate them into French, (6) *word position* (sentence-initial vs medial vs final positions), and (7) *word frequency* (absent=0 tokens, rare=1-20 tokens, frequent=21-600 tokens). Both groups were asked to listen to 20 unrelated Polish sentences recorded by a native Polish speaker and asked to repeat them as best as they could. They analysed the correct repetitions of participants concerning all the variables mentioned above. They found a significant effect of hours of instruction, that is words at eight hours of input were repeated best of all, which followed by accuracy at four hours, and words at zero hours of input were repeated least of all. Moreover, they found a strong effect of phonemic distance in the

[33] Sentences containing only 3 words were removed when effect of 8-hours of instruction was tested.
[34] In Polish, words are generally stressed on the penultimate syllable, but there are exceptions; whereas in French stress falls on the last syllable. As per design of their experiment, some words were pronounced by a Polish instructor with stress but others were unstressed.

participants' ability to repeat Polish sentences across all levels. Polish words which were phonemically close were repeated better by both groups and at all time intervals than were phonetically distant words. Likewise, the effect of lexical transparency was found across all groups. French-Polish cognates which Rast classified as very transparent were recognised best of all at all participants' levels, it followed by fairly transparent items, and the least recognised words were those which were opaque. Performance on each category of a lexical transparency variable positively correlated (increased) with the amount of exposure (from 0 to 8 hours), apart from the cognates as they were repeated well at all time-intervals (0, 4 and 8 hours). A similar effect was found for lexical stress. They found that stressed words were repeated well at all periods (0, 4 and 8 hours) unlike unstressed words although performance on them improved from 15% to 32%, and to 46% respectively with an increased amount of exposure. Additionally, they found better performance on words in the sentence-initial and final positions. Effect of frequency was found only after eight hours of exposure such that words which were frequent in the input (occurred 21-200 times) were repeated significantly better than rare words (occurred 1-20 times) and absent words (did not occur in the input). However, it needs to be noted that at zero exposure only absent in input words were compared with rare words, that is frequent words did not appear until after the testing at four hours when the frequency comparisons were not significant.

Rast (2010) compared the results of the first exposure group in the sentence repetition task to the results of the new group of first exposure learners (n=9) on the *translation task*, where they were asked to listen to the same sentences as in the sentence repetition task and were instructed to translate them into written French. She compared the results of two groups on *correct translations* and *correct repetitions* from the two tests concerning the same independent variables as above. The findings were interesting with a comparison between the sentence repetition and sentence translation showing different effects for repetition and translation. Table 3-2 below summarises the results:

	Repetitions (Period 0)	Translations (Period 0)
Word length	No	Yes
Word stress	Yes	Yes
Phonemic distance	Yes	No
Transparency	Yes	Yes
Word position	Yes	Yes
Frequency	No	No

Table 3-8. Comparison of results (adapted from Rast 2010: 75).

The table shows that the effect of word length was found for translations: longer words (3-6 syllables) were better translated than shorter (2-3 syllables). Additionally, a phonemic distance effect was important only for repetition but not for translation, but a strong effect of transparency was found for both: very transparent or cognate words were recognised significantly better than opaque and fairly transparent words. Word position for translation in both sentence-final and sentence-initial positions was important, but performance on words in sentence-final position was better. As for repetition, no such statistically significant difference was found. Transparency interacted with other variables, such as position and word stress in particular. For instance, the word *film* was not well-recognised. Why? It could be because it is a one-syllable word, making it less-salient and more difficult to perceive (see Gullberg *et al.* 2012; Carroll 2014 for discussion of the effect of syllable length) or there were other reasons.

The final study by Shoemaker and Rast (2013) was perhaps the most controlled study out of all studies by Rast herself or with colleagues. 18 native speakers of French with no previous exposure to Polish were tested at two time-intervals throughout the course: *pre-exposure* session (or zero hours of instruction) and after *6.5-hours of exposure* group session, throughout a 6.5-hour intensive Polish course on their ability to extract target words from Polish sentences as their fourth language. All participants reported L2 English and Romance languages as L3 and no knowledge of other Slavic languages, unlike Rast's previous studies (2010, 2008). The study was designed to investigate the effect of three following factors and their interaction:

1. Lexical *transparency*[35] of L2 words with respect to the L1 (high transparency vs low transparency);

2. The *frequency*[36] of the target word in the input (low frequency word = completely absent in the classroom vs high frequency = word appeared more than 20 times in the input);

3. The *position* of target words in the sentence (sentence-initial, sentence-medial vs sentence-final);

4. The number of *sessions* (0 exposure vs 6.5 hours of exposure)

A list containing 16 words in Polish was created according to transparency with respect to the L1 (French) and their frequency in the classroom. Then the list was broken into high transparency (HT) and low transparency (LT) lists. There were further broken into high frequency (HF) and low frequency (LF). There were four categories such as HT/HF, HT/LF, LT/HF and LT/LF after counterbalancing. All items comprised 2-3 syllables with the stress on the penultimate syllables. Additionally, to investigate the target word's position in the sentence, 48 test sentences were created where the target word appeared in sentence-initial, or sentence-medial, or sentence-final positions. Care was taken not to introduce a pause before or after the target words.

E-Prime software was utilised for the experiment. In each experimental trial, participants heard a sentence in Polish followed immediately by the word 'OK'. After that, they heard a Polish word in isolation and had to answer whether it had appeared in the sentence before by pressing a key on the computer keyboard. Stimuli were presented in randomised order. There was no response time limit, unlike other psycholinguistic studies on *ab-initio* learners (e.g. Gullberg *et al.* 2010, 2012; Carroll 2014).

[35] Transparency in this study was measured similarly to Rast and Dommergues (2003), by asking French native speakers with no knowledge of Slavic languages to listen to aurally presented Polish words and to translate them into French to the best of their ability. Based on results, high transparency – words with more than 50% of accuracy; low transparency – words with 0 correct translations.
[36] When selecting words for this condition, all words were counted despite their declensions.

Word recognition performance at the two sessions was compared using a repeated-measures ANOVA, where *transparency*, *frequency*, *sentences position*, and *session itself* were variables.

The significant effect of the session was found, i.e. participants' performance on the recognition of test items improved from zero exposure (accuracy=76%) to 6.5 hours (accuracy 87.9%). Concerning transparency, HT words were recognised significantly better than LT at both time intervals. However, the effect of transparency was not equal for both groups. Although, sensitivity to HT words were found for both groups, suggesting that learners may be highly dependent on phonetic and lexical forms already established in L1; and sensitivity to LT words increased significantly from the zero input session to the 6.5 hours session. Additionally, words in sentence-final position were recognised best of all at both time intervals, and better than words in sentence-initial position. Words in sentence medial position were recognised least of all in a zero exposure group, but they were recognised better than those in sentence-initial position by the second session.

Last but not least, no effect of word frequency was found, as the accuracy on HF words which participants were tested on at zero exposure (at 76.9%) was not significantly lower than accuracy on HF words after 6.5 hours of instruction (at 87.9%). Moreover, both LF (at 88.1%) and HF (at 87.9%) words were recognised equally well, with no significant difference after 6.5 hours of instruction. Shoemaker and Rast (2013) concluded that the word recognition effect of Polish words was evident after six and a half hours of instruction/exposure, but that the recognition of words does not depend on frequency (repetition of lexical items). The fact that no effect of frequency was observed is surprising, but these results should be taken cautiously because participants did not encounter any words of the frequency variable when they were tested on these words at zero amount of exposure. In other words, testing at zero amount of exposure was the very first time when participants heard LF and HF words. After that, they heard each of HF words 20 times until being tested on these LF and HF words after six and a half hours of exposure.

3.2 Summary and Predictions

To sum up, Chapter 3 clarified what is meant by the *ab-initio* learners' paradigm and objectives of this paradigm. We also discussed that using studies on artificial languages can be seen as a limitation of what should be classified as a study on *ab-initio* learners due to the fault of these studies to account for the full complexity of natural language, and what criteria the present study followed to comply with the ecological validity. It is also the reason why only studies on natural languages and those which looked at phonological processing (because of the aim of the present study) were summarised in Sections 3.1.2 and 3.1.3. These studies showed that novel words can be easily segmented from the speech for different language pairs and different methodologies employed in the *ab-initio* learners' paradigm. Additionally, these studies showed that there are several cues which facilitated the learners for segmentation. These cues are discussed next.

All studies discussed in this chapter found that an increasing amount of input positively correlated with accuracy improvement in general. However, the amount of input needed for word detection is still not very clear because it appeared to vary from study to study. For instance, Shoemaker and Rast (2013) found that Polish words were recognised well after six and a half hour of instruction to Polish. Whereas, Gullberg *et al.* (2010) found that after as little as seven minutes participants could detect monosyllabic words they encountered before in the input and they could also identify violations in phonotactics and generalise to new words of Mandarin.

The studies also showed that input could interact with other variables. For instance, Rast & Dommergues (2003) and Rast (2008) found that words which were stressed were recognised well at zero, four and eight hours of instruction/input. Also, Gullberg *et al.* (2012) showed that participants could identify in isolated forms of Mandarin bisyllabic words and frequent words (those which appeared eight times in the input) after as little as seven minutes of exposure, meaning that the recognition of these words, although improved after the quantity of this input was doubled, was not significantly better. On the contrary to the results of Gullberg *et al.* (2012), Rast and Dommergues (2003) and Rast (2008) found an effect of item frequency much later, i.e. only after eight hours of exposure to Polish in a repetition task. Similarly, Shoemaker & Rast (2013) did not find significant differences among high frequency and low frequency items at testing after six and a half hours of exposure to Polish as both words were recognised equally well. This certainly contradicts what Gullberg *et al.* (2012) found. The results of Shoemaker and Rast (2013) are caused by the fact that participants heard low and high frequency words for the first time when they were tested on these words at zero amount of input, and then they heard high frequency words 20 times before being tested again at six and a half hours. Thus, it is conceivable that when participants were tested at six and a half hours, they remembered both groups of words from the first time they were tested on them and that is why high frequency words were not recognised better than low frequency ones. Rast and colleagues concluded that these results should be taken as evidence that six and a half hours of input is sufficient for recognition of words from continuous speech, but that this recognition ability does not depend on frequency alone; instead it depends on other factors such as sentence position and transparency of target words with respect to L1 (Shoemaker & Rast 2013). Finally, Carroll's (2014) study showed that to learn cognates required less training trials than to learn non-cognates and that beginners in German required fewer trials than a first exposure group. Also, there were individual differences such as that some participants learned all target names just with two trials, but others needed eight trials.

As a matter of fact, some studies showed that *no* input at all or very little input is needed for participants to show sensitivities to forms of target words. Gullberg *et al.* (2010) found that no exposure to Mandarin was sufficient for Dutch L1 participants who heard syllables which violated Mandarin syllables structure (Gullberg *et al.* 2010) to recognise them as not-Chinese. These findings are perhaps not surprising provided that exposure to Chinese is widespread, recall that according to the Office of National Statistics, 1% of the population within English and Wales speak Chinese as their main language (Potter-Collins 2013). If it happens in the UK, it probably happens in other parts of the world with around 16 % of the world population speaking Chinese as their first language. Additionally, with zero exposure, French L1 participants could accurately translate and repeat Polish words which were very transparent (cognate) with French (Rast & Dommergues 2003, Rast 2008, 2010) and to recognise these words in a sequential context (Shoemaker & Rast 2013). These findings are consistent with Carroll (2014) who also found that cognate names were recognised well after the first few trials of exposure even by first exposure learners (with no exposure to German). These results demonstrate that *ab-initio* learners can detect target words if they share similarities with the ones in their native language, which provides evidence that L1 transfer operates at the very beginning of L2 development. Moreover, it is clear that repeated exposure to some aspects of language may not be necessary, and sensitivity to some aspects of language may require no exposure at all. The hypotheses of this study are formulated after the methodology is discussed in Chapter 5. However, based on the findings discussed above, it can be predicted that participants' accuracy on words from the input will increase with an increasing amount of sessions. Additionally, given the robust effect of cognates or phonologically transparent items, it can be expected that participants performance on cognate words would be high from the very first moment of encountering these words[37].

Going back to the discussion of findings from *ab-initio* learners, the length of words cue was included in the analyses of every study discussed above. All studies found evidence that longer words are more salient for learners than shorter words. Bisyllabic words in Gullberg *et al.* (2012) were segmented better than monosyllabic words, and names with up to seven syllables were segmented better than shorter names in Carroll (2014). Additionally, Rast (2010) discovered that longer words with 3 to 6 syllables were better translated than shorter words from 0-1 syllables and 2 syllables, but Rast and Dommergues (2003) and Rast (2008) did not find the effect of length of words in the sentence repetition task either at zero or after four or eight hours of input. Although participants' were more accurate in repeating shorter words (0-1 syllables) than longer words (3-6

[37] The present study utilised cognate identification task in order to measure participants' ability to pay attention to the input. It is discussed in Section 5.3.1.3.

syllable) in the sentence-initial position, and there was an opposite effect for the interaction of sentence-final position and longer words so that longer words were repeated better than shorter words. Rast (2008) suggested that repetition of words could somehow depend on the interaction of word length and sentence position, but she did not elaborate any further. This effect could be attributed to production. As the present study is on perception, it is reasonable to expect that longer words would have higher success rates than shorter words when they are detected from a continuous speech stream. This is consistent with studies on the effect of distributional properties on segmentation in artificial languages (e.g. Saffran *et al.* 1996b) described in 2.1.3, in a way that words which consist of more than one syllable have higher word-internal transitional probabilities which are easier to compute in comparison, for instance, with monosyllabic words which have only word-external transitional probabilities. Therefore, it is easier to detect words of more than one syllable in a speech stream. With respect to the present study, it can be predicted that learners would be more accurate in detecting Russian bisyllabic words than Russian monosyllabic words. Moreover, thinking about an effect of input as discussed above, it can be predicted that participants of this study will be more accurate in detecting Russian bisyllabic words than Russian monosyllabic words, and this ability will increase over sessions.

Concerning the sentence position of a target word, it was generally found that words at the sentence-final position were segmented best throughout all sessions, which was followed by the sentence-initial position (Rast 2010). Similarly, Carroll (2014) found that words were equally segmented in both sentence-medial and -final positions. As you will see from Chapter 5, sentence position was not a variable which was investigated in the present study; instead the effect of sentence position was kept constant.

It was mentioned at the beginning of this chapter that generalisation beyond exemplars in the input represents the learner's problem of L2 analysis. Among all studies on *ab-initio* learners, only one study by Gullberg *et al.* (2010) showed that learners could generalise phonotactic information from Mandarin Chinese to novel items not encountered in the input after as little as a seven minute of exposure to Mandarin Chinese. *Generalisation* is usually defined as a transfer of prior learning to new situations and problems (Gluck *et al.* 2008: 337). According to Gluck *et al.* (2008) psychologists have studied extensively the generalisation of learning, but it is still not clear how learning one thing can be generalised to another, and why some generalisations have limits. With respect to the generalisation in language learning, most of the research which has been done with respect to natural language is on the acquisition of morphosyntax (see Prasada & Pinker 1993; Christiansen & Chater 1994; Goldberg 2006). However, research on generalisation in the phonological aspects of natural language is very scarce, with the exception of a couple of studies on artificial languages which focused on the end stage of language knowledge, e.g. Finley and Badecker (2009); Cristia *et al.* (2013) and only one study on the early stage by Linzen and Gallagher (2017). Drawing inspiration from the study on natural language learning by Gullberg *et al.* (2010), the present study is going to investigate if *ab-initio* learners of Russian can generalise to novel stimuli after exposure to Russian. Based on findings from Gullberg *et al.* (2010), it can be predicted that *ab-initio* learners of Russian will have generalisation ability after minimal exposure to Russian. Finally, but importantly, it has been shown that children (18-24 months) learn words with one exposure (Bloom 2000), in L2 studies it is generally recognised that although adults are competent vocabulary learners, they are rarely granted the capacity for fast mapping. The results from Gullberg *et al.* (2010, 2012) study does not support this as they showed that participants were capable of fast mapping of the items they encountered in the input after just 7-minutes of watching a clip where word-referent mapping was facilitated by pointing gestures. Additionally, Carroll (2014) found similar results as her *ab-initio* and beginners in German groups could segment cognate and compound name and match them to people-referents from the pictures after only a few trials of exposure. This is fascinating as it clearly shows that *ab-initio* learners can fast-map words to meanings. It is undoubtedly important as it is one of the steps in the learner's problem of analysis (1986). However, as already mentioned several times, this study is not concerned with how learners establish meaning. Instead, the present study focuses on segmentation and generalisation ability from exposure to novel items. The next Chapter 4 reviews the language background of English and Russian languages.

Chapter 4. Language background

4.1 Introduction

This chapter focuses on a discussion of English and Russian phonologies. It has already been mentioned in Section 2.2.5 that there are language-specific phonological constraints in each language, which means that, for example, what is possible or allowed in Russian is not necessarily allowed in English, and what is legitimate in English may be illegitimate in Russian. Moreover, main notions such as SSP and MSD were also discussed. We already know that SSP refers to a universal principle which assigns structure to syllables in terms of the sonority of its segments and that MSD is a language-specific realisation of segments within a syllable. In addition, it was mentioned in Section 2.2.4 that stress refers to the prominence of one syllable over another in a sequence of syllables. The stressed syllable involves more muscular effort in its production; it is louder, longer and shows more pitch variation than the surrounding syllables (Davenport & Hannahs 2010: 78). This chapter demonstrates that both English and Russian are Indo-European languages, and there are similarities in phonologies of these two languages; for example, many sounds, phonological processes and phonotactic constructions which are possible in English are possible in Russian as well. There is also similarity in these languages in the way stress is realised by reducing non-high vowels in unstressed syllables. However, this chapter also demonstrates that there are certainly differences in the phonetic inventory, phonotactic constraints and the mechanisms of stress assignment between Russian and English. The chapter starts with a discussion of English in Section 4.2, which is followed by a discussion of Russian in Section 4.3. Both begin with an overview of the phonetic inventory, and the most important phonological processes, followed by a discussion of phonotactics and stress. A summary of the chapter and its predictions is given in Section 4.4.

4.2 The English language

The tables below illustrate most of the consonants, vowels and diphthongs, which the varieties of the English language typically have. There are variations in terms of the consonant sounds English varieties allow, but there is even more variation with respect to vowel sounds (see Davenport & Hannahs 2010). Table 4-1 below shows that there are 24 consonants, with the voiceless velar fricative /x/ represented as optional because nowadays it is found in Scottish English (e.g. 'broch' [brɒx]), and in Welsh (e.g. 'dear' [bɑːx]). Table 4-2 shows that there are 12 vowel sounds and eight diphthongs based on Standard British pronunciation, but as already said this may vary depending on the variety.

		Labial		Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
		Bi-labial	Labio-dental						
Stop	-voice	p			t			k	
	+voice	b			d			g	
Affricate	-voice					tʃ			
	+voice					dʒ			
Fricative	-voice		f	θ	s		ʃ	(x)	h
	+voice		v	ð	z		ʒ		
Appro-ximant	-voice								
	+voice	w			ɹ		j		
Nasal	-voice								
	+voice	m			n			ŋ	
Lateral	-voice								
	+voice			l					
Trill	-voice								
	+voice								
Tap / Flap	-voice								
	+voice								

Table 4-1. Phonemic inventory of English consonants.

	Front		Central		Back	
	short	long	short	long	short	long
High	I	iː			ʊ	uː
Mid	e (ɛ)		ə	ɜː		ɔː
Low	æ (a)		ʌ		ɒ	ɑː
Diphthongs	eɪ, aɪ, ɔɪ, aʊ, əʊ, ɪə, eə, and ʊə					

Table 4-2. Phonemic inventory of English vowels and diphthongs based on Standard British pronunciation.

Table 4-1 shows that most obstruents of English come in voiced-voiceless pairs, except for /h/. One of the unique phonological features of English is that voiceless stops can be aspirated when found in the word-initial position, as discussed in Chapter 2. Moreover, English is unlike many European languages in that it does not have word-final devoicing, which means that many words in English maintain a voicing contrast in the word-final position such as ‘cap’ [kap] vs ‘cab’ [kab], or ‘seat’ [siːt] vs ‘seed’ [siːd]; but see Docherty (1992) for a discussion of a trend towards partial or whole word-final devoicing which appears to exist in English and which makes English and Russian similar in this respect[38]. Moreover, English and Russian are similar in terms of voicing assimilation (Russian word-final devoicing and assimilation is discussed in detail in Section 4.3) but, unlike Russian assimilation, English assimilation is morphologically conditioned. This means that any time an inflectional allomorph {-s} (whether plural or possessive), which is a voiceless alveolar fricative, is added to a word containing a voiced segment at the end, an allomorph becomes voiced (e.g. ‘dog’ [dɒg] → ‘dogs’ [dɒgz]); but if it is attached to a word ending on a voiceless segment, the /s/ remains voiceless (e.g. ‘cat’ [kat] → ‘cats’ [kats]). Note that this is not a phonetically conditioned assimilation, as English phonology does allow voiced segments to be followed by voiceless segments (e.g. ‘fence’ [fɛns]) just as Russian does (e.g. *seans* [sʲɪˈans] ‘session’). Finally, English consonants can be assimilated according to the place of articulation; for example, the alveolar nasal /n/ can become the bilabial nasal /m/ before bilabial stops (e.g. ‘input’ [ˈɪmpʊt]).

4.2.1 English phonotactics

English allows syllables without an onset and a coda, and it also allows both word-initial and word-final consonant clusters, just as branching nuclei are legal as well. An English nucleus can contain all vowel and diphthong sounds, as well as /l/, /m/ and /n/ sounds, and /r/ sounds, but this

[38] The phonology of Russian is discussed in Section 4.3.

is exclusive to rhotic varieties of English[39]. The English onset allows between zero and three consonants in its position. An English CV syllable can start with any consonant, with the exception of [ŋ], which can never appear in a syllable-initial position. Table 4-3 below demonstrates allowable onset types.

	n of segments	example	transcription
1.	zero segment	‘eye’	[ʌɪ]
2.	one segment	‘buy’	[bʌɪ]
3.	two segments	‘smile’	[smʌɪl]
4.	three segments	‘sprout’	[sprəʊt]

Table 4-3. Allowable onset types in English.

English clusters with two consonants must be rising in sonority. If the first segment in a cluster is an obstruent, it should be followed by either a liquid /l, r/, (e.g. ‘play’ [pleɪ]) or glide /w, j/ (e.g. ‘twin’ [twɪn]); but if the first segment in a cluster is a voiceless fricative /s/, then it should be followed by either nasals, but not except [ŋ] (e.g. ‘snow’ [snəʊ]) or voiceless stops (e.g. ‘speak’ [spiːk]), or approximants (e.g. ‘slope’ [sləʊp]). This means that the MSD of English clusters must be equal to two. Harris (1994) states that such English clusters with MSD=2 are common to many languages in the world, and therefore are highly permitted universally.

Three-member onset consonant clusters in English must follow the specific rule that the first consonant in such a cluster must be /s/, which should be followed by a voiceless oral stop and a liquid or a glide (e.g. ‘spring’ [sprɪŋ], ‘split’ [splɪt], or ‘stew’ [stjuː]). Unlike in Russian, which is discussed next, English allows neither sonority plateaus nor reverse sonority.

An English coda can contain between zero and four consonants. Any consonant can appear in a syllable-final position, except for /h/, examples of which are presented in Table 4-4 below.

[39] These sounds act as syllabic consonants, for example in words like ‘even’ [ˈiːvn̩], ‘little’ [ˈlɪtl], where the /n/ and the /l/ occupy the nucleus position of a second syllable.

	n of segments	example	transcription
1.	zero segment	‘see’	[siː]
2.	one segment	‘seek’	[si:k]
3.	two segments	‘six’	[sɪks]
4.	three segments	‘sixth’	[sɪksθ]
5.	four segments	‘sixths’	[sɪksθs]

Table 4-4. Allowable coda types in English.

4.2.2 English stress

The main stress placement in English has been extensively discussed by linguists. There is widespread agreement that English stress can be predicted from the phonological properties of a word, where a lexical class and syllable structure determine stress placement. However, stress placement in English can also be lexical, which means that the position of stress in certain words cannot be predicted by a rule, and so it needs to be memorised. Moreover, in morphologically complex words, stress placement is morphologically conditioned. As this thesis does not aim to provide a full explanation of stress placement in English, this section focuses on explaining how phonological factors predict stress in English with the examples of underived nouns and verbs. In fact, it would be sufficient to look at just English nouns, because the aim of this thesis is to investigate how Russian bisyllabic nouns are detected by the means of stress in a continuous speech stream, and so, by hypothesis, English L1 learners should tap into their knowledge of English nouns but they may in fact also tap into their knowledge of English verbs and even adjectives. Moreover, explaining how stress functions in English verbs in addition to English nouns may provide a much complete picture of stress assignment in English. However, before stating the details of English stress placement in nouns and verbs, it needs to be mentioned that English is essentially a quantity-sensitive language[40], which means that heavy syllables attract stress, which are those having either a long vowel in a nucleus (e.g. CVV) or closed syllables with at least one consonant in a coda (e.g. CVC); as opposed to light syllables – those which have a short vowel in a nucleus and do not have a coda (e.g. CV), which are usually not stressed.

[40] There are quantity-insensitive languages, such as Polish, Hungarian and French, in which the position of stress in these languages is fixed, and syllable weight has no importance in stress assignment.

In the case of nouns, English exhibits *extrametricality*, which means that a final syllable of a word can be dropped from considering in terms of being stressed, and the stress then moves to the penultimate syllable, and if in turn, this syllable is heavy, it receives stress (e.g. ‘potato’ [pəˈteɪtəʊ], or ‘adventure’ [ədˈvɛntʃə]; but if the penultimate syllable is not heavy, the stress moves to the antepenultimate syllable (e.g. ‘family’ [ˈfamɪli]). What this means is that most bisyllabic nouns would be stressed on the first syllable; except for those which have a long vowel in their final syllable where stress then falls on the last syllable (e.g. ‘guitar’ [gɪˈtɑː], and ‘surprise’ [səˈprʌɪz]).

In the case of two-syllable verbs, the final consonant of a word is treated as extrametrical; that is, it is excluded from a possibility being stressed and the stress placement starts from what is left in the final syllable of that word. Just like with deciding on the weight of penultimate syllables in nouns, English verbs are approached in the same way. That is, if the last syllable (without an extrametrical final consonant) is heavy, it receives stress (e.g. ‘record’ [rɪˈkɔːd], ‘reveal’ [rɪˈviːl]); and if it is not, the stress moves to the next syllable on the left (e.g. ‘exhibit’ [ɪgˈzɪbɪt]).

It appears from the above that, in English, bisyllabic nouns are more likely to have the main stress on the first syllable, whereas bisyllabic verbs are more likely to be stressed on the final syllable. Furthermore, there is evidence from Sereno, and Kelly and Block 1988 (1986 and 1988 cited in Guion *et al.* 2003: 406) that bisyllabic nouns can be stressed on the first syllable (73% of the time in Sereno, and 94% of all times in Kelly and Block. On the other hand, English verbs were stressed on the first syllable 34% of the time by Sereno (1986) and 31% by Kelly and Block (1988). These figures were arrived at by looking at stress assignment in two-syllabic nouns and verbs using the frequency corpus of Francis and Kučera (1982). Finally, as described in Section 2.2.4, about 90% of all lexical words in English are stressed on the first syllable, and word-initial stressed syllables can be used as a cue in the segmentation of continuous speech by assuming that every new word starts at a strong syllable (Cutler & Carter 1987).

4.3 The Russian language

There are 37 consonants and six vowels in the Russian language. Table 4-5 below, adapted from Timberlake (2004), summarises its phonemic inventory. The table shows that one of the distinguishing properties of Russian is contrastive palatalisation, which means that certain consonantal sounds have a pair (i.e. palatalised/plain); for instance, *mat* [mat] ‘checkmate’ vs *mat’* [mat^j] ‘mother’. Table 4-5 illustrates that most consonants in Russian have a palatalised counterpart. It needs to be mentioned that there are certain properties of the Russian phonetic inventory, which are viewed as controversial by researchers, see Chew (2000) and Timberlake (2004), for example, for discussions. Among these issues is the status of the high central unrounded vowel [ɨ]. Some linguists believe that [ɨ] is an allophone of /i/; whereas others assume that /i/ and /ɨ/ are independent phonemes. Moreover, some phonetic inventories of Russian include the phonemes /k^j, g^j, x^j/, whereas others do not because certain linguists believe that the Russian velars [k^j, g^j and x^j] are allophones of /k, g, x/ respectively. Finally, the status of long palatalised fricatives [ʃ^j:] and [ʒ^j:] as independent phonemes has been discussed. The arguments underlying each point of view are not relevant to this study, but these controversial phonemes are indicated in parentheses in the tables below.

		Labial				Dental		Alveo-palatal		Velar	
		Bilabial		Labio-dental							
		Plain	Palat.	Plain	Palat.	Plain	Palat.	Plain	Palat.	Plain	Palat.
Stop	-voice	p	p^j			t	t^j			k	(k^j)
	+voice	b	b^j			d	d^j			g	(g^j)
Affricate	-voice					ts			tʃ^j		
	+voice										
Fricative	-voice			f	f^j	s	s^j	ʃ	(ʃ^j:)	x	(x^j)
	+voice			v	v^j	z	z^j	ʒ	(ʒ^j:)		
Approximant	-voice										
	+voice								J		
Nasal	-voice										
	+voice	m	m^j			n	n^j				
Lateral	-voice										
	+voice					l	l^j				
Trill	-voice										
	+voice							r	r’		

Table 4-5. Phonemic inventory of Russian consonants: Palat.=palatalised.

	Front	Central	Back
High	i		u
Mid	ɛ (ə)	(ɨ)	o
Low		a	

Table 4-6. Phonemic inventory of Russian vowels.

The Russian phonetic inventory has both voiced and voiceless obstruents, and it is important to take note of two phonological processes that involve voice features. Firstly, Russian obstruents are subject to voicing assimilation, which applies within words and across word boundaries. That is, if there are sequences of obstruents within a word, they must agree via a voicing feature with the last segment within this sequence (e.g. *lodka* [ˈlotkə] ‘boat’) where the voiced /d/ segment becomes voiceless; or vice versa (e.g. *skazka* [ˈskaskə] ‘fairy-tale’) where a voiceless segment becomes voiced. Additionally, if there are sequences of obstruents across words’, they also must agree in terms of voice (e.g. *pod solntsem* [ˈpətˈsontsɨm] ‘under the sun’, and *ot doma* [ˈod ˈdomə] ‘from the house’). However, voicing assimilation applies with the exception of sonorants in Russian which are voiced by default; for example, /r/, /l/, /m/, /n/ and /j/. That is, when any voiceless consonant is followed by any of these sonorant sounds, it does not become voiced (e.g. *sestra* [sʲɪˈstra] ‘sister’, *kniga* [ˈknʲigə] ‘book’). Moreover, the voiceless sounds /t/, /k/ and /ts/ are not affected by voiced sounds preceding them (e.g. *kvas* [kvas] ‘kvass’). Secondly, Russian voiced obstruents become voiceless at the end of a word (e.g. *prud* [prut] ‘pond’, and *ogorod* [agaˈrot] ‘garden’).

4.3.1 Russian phonotactics

Any sound in the Russian phonetic inventory can start and end a word in Russian, perhaps with the single exception that the mid-central vowel /ɨ/ cannot start a word. One of the striking facts about Russian phonology is the variety of consonant clusters which are allowed in the language. In Russian, both word-initial and word-final consonant clusters are allowed, and no branching nucleus is permitted. According to one of the first descriptions of Russian phonology (Halle 1959), a Russian onset can contain between zero and four positions. For instance:

	n of segments	Russian	transcription	translation
1.	zero segment	*um*	[um]	‘mind’
2.	one segment	*dom*	[dom]	‘house’
3.	two segments	*dl’ia*	[dlʲa]	‘for’
4.	three segments	*skrip*	[skrʲip]	‘squeak’
5.	four segments	*vstretit’*	[ˈfstrʲetʲɪtʲ]	‘to meet’

Table 4-7. Russian syllable onset types.

As for clusters with two segments, Russian allows many clusters with rising sonority, or MSD=2 (e.g., *sv’et* [svʲet] ‘light’), and MSD=1 (e.g., *sn’eg* [sʲnʲek] ‘snow’). Moreover, it allows clusters with plateau sonority, or MSD=0, (e.g., *kniga* [ˈknʲigə] ‘book’). Finally, Russian allows clusters even with reverse sonority or MSD=-1 (e.g., *rta* [rta] ‘GEN. SG. mouth’) which, however, are rather rare based on the corpus frequencies from Sharoff (2002 cited in Proctor 2009: 129-132).

As Table 4-7 shows, Russian allows three-segments clusters in the onset, but their sequence combinations are specific to the following rule as described in Chew (2000) and Trapman (2007):

a) [f] or [v] + [s] or [z] + sonorant, e.g. *vzr’yv* [vzrɨf] ‘explosion’;
b) [f] or [v] + [s] or [z] + stop, e.g. *vskor’e* [ˈfskorʲɪ] ‘soon’;
c) [f], [v] or [s], [z] + stop + liquid or /v/, e.g. *zdravstv’yjt’e* [ˈzdrastvʊjtʲe] ‘hello’.

Four segment clusters are also specific to the rule in that they all must begin with /fs/ or /vz/, with regressive voicing assimilation. Trapman (2007) points out that the majority of CCCC onsets have a stop in the third and a liquid in the fourth positions in the onset, such as in *vzbros* [vzbros] ‘upthrust’ or *vstrecha* [ˈfstrʲetʃ:ə] ‘meeting’.

Just as with the Russian onset, a Russian coda can also contain between zero and four positions. Although codas in Russian do not represent a particular interest in this study, examples with all possible coda types in Russian are given in Table 4-8 below.

	n of segments	Russian	transcription	Translation
1.	zero segment	*n'u*	[nu]	'(colloquial) yeah,yep'
2.	one segment	*dom*	[dom]	'house'
3.	two segments	*kost'*	[kosʲtʲ]	'bone
4.	three segments	*tolst*	[tolst]	'MASC. SG. fat'
5.	four segments	*ch'orstv*	[tʃ:ərstf]	'stale'

Table 4-8. Russian coda types.

4.3.2 Russian stress

Russian stress assignment resembles those features which are found in English and many European languages in the sense that each lexical word has one syllable which bears primary stress. Secondary stress is possible in Russian as well, but it is restricted to compound words. Like in English, Russian stressed syllables have greater duration and higher pitch, which leads them to be perceived as louder and longer, whereas non-stressed syllables are reduced making them perceived as less prominent (Jones & Ward 1969).

In this section, the stress patterns only of Russian nouns is discussed. This is because, firstly, most research has been conducted on Russian nouns, and secondly because the present study looks at how Russian monosyllabic and bisyllabic nouns are detected in a continuous speech stream.

There have been many attempts to describe the factors underlying stress patterns in Russian. Some linguists (e.g. Zaliznjak 1977, 1985; Archibald 1994) assume that Russian belongs to those languages with *unpredictable* or '*free*' stress, which means that stress must be stored as part of the lexicon and that the phonological properties of a word do not influence stress assignment (unlike in English, as was illustrated in Section 4.2.2, where syllable weight influences stress assignment). Meanwhile, other researchers (e.g. Halle 1975, 1997; Melvold 1990; Alderete 2001; Crosswhite *et al.* 2003) believe that there must be some underlying phonological and morphological principles which govern stress assignment in Russian.

In my understanding, the most comprehensive overview of stress assignment in Russian nouns is provided by Alderete (2001). According to him, there are three ways stress can be positioned in underived Russian nouns. Before moving to a discussion of these, it is important to mention that underived Russian nouns consist of a stem which is followed by a grammatical inflection which specifies gender, number and case. There are three grammatical genders: masculine, feminine and neuter. Also, number can be singular or plural. Besides this, there are six cases: nominative, accusative, genitive, dative, instrumental and prepositional.

Firstly, there is a fixed-stress pattern on a stem, which means that any syllable within a stem of a word can be stressed, but when this word form becomes declined, none of the inflections are stressed. This is illustrated below by an example from the word *kniga* [ˈknʲigə] ‘book’, which is a feminine noun declined by number and case. It is seen from the paradigm in Table 4-9 that the stem remains stressed across all instances; that is, the stress never shifts to the ending when a word is declined.

Singular number			Plural number		
case	Russian	transcription	case	Russian	Transcription
Nominative	*kn'ig-a*	[ˈknʲig-ə]	Nominative	*kn'ig'-i*	[ˈknʲigʲ-ɪ]
Accusative	*kn'ig-y*	[ˈknʲig-ʊ]	Accusative	*kn'ig'-i*	[ˈknʲigʲ-ɪ]
Genitive	*kn'ig'-i*	[ˈknʲigʲ-ɪ]	Genitive	*kn'ig*	[ˈknʲig]
Dative	*kn'ig'-e*	[ˈknʲigʲ-e]	Dative	*kn'ig-am*	[ˈknʲig-əm]
Instrumental	*kn'ig'-e*	[ˈknʲigʲ-e]	Instrumental	*kn'ig-ami*	[ˈknʲig-əmʲɪ]
Prepositional	*kn'ig'-e*	[ˈknʲigʲ-e]	Prepositional	*kn'ig-ah*	[ˈknʲig-əx]

Table 4-9. Paradigm with fixed stress pattern on a stem in Russian.

Secondly, there a fixed-stress pattern on an inflection, which means that when a word becomes inflected, the stress from a stem shifts to an inflection and any syllable within the inflection can be stressed, whereas the stem remains without stress. This stress placement is illustrated by an example of the word *zamok* [zaˈmok] ‘lock’, which is a default, masculine form, in Table 4-10 below. The paradigm below demonstrates that when this word declines for number and gender, the stress always shifts from the last syllable of a stem to inflectional endings.

Singular number			Plural number		
Case	Russian	transcription	case	Russian	transcription
Nominative	*zamok*	[za'mok]	Nominative	*zamk-u*	[zam'kʲ-i]
Accusative	*zamok*	[za'mok][41]	Accusative	*zamk-u*	[zam'kʲ-i]
Genitive	*zamk-a*	[zam'k-a]	Genitive	*zamk-ov*	[zam'k-of]
Dative	*zamk-y*	[zam'k-u]	Dative	*zamk-am*	[zam'k-am]
Instrumental	*zamk-om*	[zam'k-om]	Instrumental	*zamk-ami*	[zam'k-amʲi]
Prepositional	*zamk-e*	[zam'kʲ-e]	Prepositional	*zamk-am*	[zam'k-ax]

Table 4-10. Paradigm with fixed-stress pattern on inflection in Russian.

Thirdly, there is a mobile-stress pattern in Russian, which means that in the given word paradigm, some forms of a word receive stress on a stem, while other word forms receive stress on inflections. This stress pattern is illustrated in Table 4-11 with declensions of the word *delo* ['dʲelə] 'business'. The paradigm shows that stress always falls on the stem in the singular form across all cases, but whenever inflectional endings are added to the stem in the plural, the stress shifts from the stem to inflectional morphemes in some cases but not in others.

Singular			Plural		
case	Russian	transcription	case	Russian	transcription
Nominative	del-o	['dʲel-ə]	Nominative	del-a	[dʲɪ'l-a]
Accusative	del-o	['dʲel-ə]	Accusative	del	['dʲel]
Genitive	del	['dʲel]	Genitive	del-am	[dʲɪ'l-am]
Dative	del-u	['dʲel-ʊ]	Dative	del-a	[dʲɪ'l-a]
Instrumental	del-om	['dʲel-əm]	Instrumental	del-ami	[dʲɪ'l-amʲɪ]
Prepositional	del-e	['dʲelʲ-e]	Prepositional	del-ah	[dʲɪ'l-ax]

Table 4-11. Paradigm with mobile stress pattern in Russian.

It follows from the above that stress in Russian can fall on any vowel within a stem, or the first vowel of the inflectional ending. Corpus frequency counts by Zaliznjak (1977) showed that words with the fixed-stress pattern on a stem constitute 92% of the total, whereas words with the fixed-stress pattern on inflection constitute 6%, and the remaining 2% represent words with the mobile-stress pattern.

Although this descriptive account explains some facts about stress assignment in Russian nouns, it still not clear why some words follow the fixed-stress pattern on a stem, while others follow the fixed-stress pattern on inflection or a mobile stress pattern. Also, unlike in English, the underlying

[41] An accusative form of masculine, singular nouns in Russian is the same as a default form or nominative case.

principles in Russian which explain stress placement in words which consist of more than one syllable are not clear. The likely reason for the lack of a clear account which explains the underlying principles of Russian stress placement is because there is so much variability in how Russian L1 speakers pronounce stress, as well as among Belarusian-Russian, Kazakhstan-Russian or Ukrainian-Russian speakers. Many of these speakers acquire Russian *via* reading but in books, words are not accented unless they are intended for children or have been adjusted to suit a foreign reader. Hence, when these learners encounter a new word, many guess at the stress placement instead of checking in a dictionary for the correct pronunciation, which leads them to learn words with incorrect stress placement. Nearly all polysyllabic words in Russian can be stressed incorrectly (Lebedeva 1986), for example: *oblegchit'* 'to ease' is often pronounced as [ɐbˈlʲexˈtʃʲitʲ] instead of [ɐblʲɪxˈtʃʲitʲ], *zvonit'* 'to call' is often pronounced as [zvˈonʲitʲ] instead of [zvɐˈnʲitʲ] or *tort* 'cake' is pronounced as [ˈtɐrˈtɨ] instead of [ˈtortɨ], the former versions of which are considered correct pronunciations in Standard Russian. So, if correct stress placement is difficult for native speakers themselves, it would obviously be difficult for L2 learners of Russian. Hart (1998) showed in a production study where passages in Russian needed to be read that 06correct stress placement was difficult even for advanced L1 English learners, and the stress was misplaced in one word out of five. To sum up, it must be evident from the above that stress placement in Russian is complex, and this section cannot provide a complete picture of stress placement in Russian nouns[42]. However, the information provided in this chapter on stress placement in Russian and in English should be sufficient for the formulation of hypotheses for the present study.

[42] A detailed understanding of Russian stress placement can be gained from Zaliznjak (1977, 1985); Melvold (1990); Alderete (2001); Crosswhite *et al.* (2003); Lavitskaya (2015).

4.4 Summary and predictions

Chapter 4 has provided an overview of relevant aspects of the phonologies of the English and Russian languages. Observations concerning similarities and differences between these two languages are discussed here and, as the target language in this study is Russian, the emphasis is placed on what is found in Russian but does not exist in English. Firstly, most consonantal sounds in the phonetic inventories are the same, with the exception that many consonants in Russian have palatalised counterparts. Moreover, Russian lacks interdental fricatives, but it has pairs of two sounds such as /ʃʲː/ and /ʒʲː/, as well as /x/ & /xʲ/, and an affricate /ts/ which are not found in English. Besides this, although both languages allow rhotic sounds, the Russian rhotic is an alveolar palatal trill /r/, while most varieties of English have an alveolar approximant /ɹ/. Most importantly, English is a superset of Russian with respect to the vowel sounds it allows, as there are only six vowel sounds in Russian, where the mid-central vowel /ɨ/ is specific to Russian. Also any vowel sequence in Russian is treated as a vowel hiatus by most researchers, but English allows varieties of diphthongs.

In terms of phonological processes, it must be mentioned that Russian is different from English in that all voiced stops in a word-final position are devoiced. English has a word-final voicing contrast, although some linguists believe that word-final devoicing is also common cross-linguistically, and moreover occurs in English too; perhaps more in some dialects than others. Furthermore, both languages appear to manifest voicing assimilation, where two consonant sounds next to each other agree in terms of the voice feature. Where they differ is that English has progressive voicing assimilation, the application of which is morpho-phonologically determined, but Russian has a phonologically determined regressive voicing assimilation. It needs to be mentioned here that there are several exceptions to the rules of voicing assimilation in both languages.

As for phonotactics, all sounds in Russian can appear in a word- or syllable-initial position, as well as in word- or syllable-final positions with the exception of /ɨ/, but no English syllable can start with /ŋ/. This is unlikely to be relevant in any case as the Russian phonetic inventory does not have a velar nasal. However, the /h/ sound in a word-final position is not allowed in English. Again, as /h/- the glottal fricative sound is absent from the phonetic inventory, and so we do not need to worry about it. However, there is another fricative in Russian, which is the velar /x/, and some words in Russian do end with this segment. Consequently, as /xʲ/ and /h/ have the same manner of articulation and although different are very close in terms of the place of articulation, it could potentially be seen as troublesome for L1 English learners that syllables and words in Russian can have /x/ as their final segment. Moving to consonant clusters, it must be said that languages allow varying combinations of MSD values. For example, Russian allows MSD=0 because such clusters as [pt] or [kt] are both obstruents are attested in the language, and Russian can even flout sonority with MSD= -1. On the other hand, MSD in English can drop below a value of 2, e.g. [pl], [kr], [tw].

To sum up the above paragraphs concerning the phonetic inventories and phonotactics of English and Russian, it can be suggested that L1 English learners of L2 Russian would encounter only a few unfamiliar segments (such as /ʃʲ:/, /ʒʲ:/, /x/ & /xʲ/, /ts/, and /ɨ/) in Russian and, along with a comparatively small number of vowels and a complete absence of diphthongs in Russian, the phonetic inventory of Russian should not represent a major challenge for L1 English learners. However, detecting words in Russian, which start with patterns of phonotactics which L1 English learners do not have in their L1 will undoubtfully cause difficulties. It can be predicted that participants in the present study would be more accurate in detecting patterns of Russian phonotactics which exist in English. Additionally, it was mentioned in Section 2.1.3 that infants and adults could use distributional cues for the detection of words in a speech stream. Also, the results of studies on *ab-initio* learners show that adults can segment longer words better than shorter words. Based on these phenomena, we formulated a prediction in Section 3.2, that is

Russian bisyllabic words would be segmented better than monosyllabic words. As one of the aims of the present study is to investigate the effect of the interaction of cues, it can also be predicted that participants in this study will be more accurate in detecting words which phonotactics occur in both English and Russian, and that there should be a preference for bisyllabic words.

Concerning stress placement, it should be clear that English stress can be predicted from the phonological properties of a word, where word class and syllable stress determine stress placement, whereas Russian stress placement cannot be predicted according to an underlying rule. Despite the differences in stress placement in these two languages, there is an interesting observation in terms of how English and Russian treat borrowed words from each other's language. Hart (1998) notes that when in the Russian language a word is borrowed from English, English stress is retained, and the stress pattern of English is not changed into the Russian form; e.g. *proˈfesor* [praˈfʲesər] in Russian was borrowed from English 'professor' [prəˈfɛsə]; and *doctor* [ˈdoktər] was borrowed from 'doctor' [ˈdɒktə]). On the contrary, when Russian nouns are borrowed into English, their stress is adjusted following the pattern discussed in Section 4.2.2; e.g. 'babushka' [bəˈbʊʃkə] but the Russian is *babushka* [ˈbabuʃkə], and 'gulag' [ˈguːlɑːg] in English but the Russian is *gulag* [gʊˈlak]. This surprising fact leads us to believe even more strongly that when English learners are confronted with Russian, they are confronted with idiosyncratic stress assignment, which Hart (1998) refers to as a system which is unrelated to their native one. Moreover, the fact that English adjusts the stress patterns of borrowed words is likely to prove the existence of an underlying stress system in English, and the flexibility found in Russian suggests that there is no underlying pattern.

As there is no clear pattern (or underlying phonological principle) for how stress is placed in Russian nouns, no clear prediction can be made about what to expect in terms of Russian stress placement, and thus it was hard to make any predictions that English L1 learners would show sensitivity to a particular strategy of stress placement in Russian bisyllabic nouns. However, if L1

English learners bring their knowledge of MSS (as discussed in Section 2.1.4) into the analysis of Russian input, it can be predicted that they should be better in detecting Russian bisyllabic words which are stressed on the first syllable rather than on the second syllable. Additionally, as discussed in Section 2.1.6, it is clear that phonotactics and stress can interact in bisyllabic words, so it can be predicted that participants in this study will be more accurate in detecting Russian bisyllabic words which have phonotactics which are held in common by English and Russian and which are stressed on the first syllable.

Finally, it was discussed in Chapter 3 that all studies on *ab-initio* learners have found that an increasing amount of input positively correlates with improvements in accuracy in general. As the present study investigates the effect of the interaction of segmentation cues with input, based on what was discussed above, the following two predictions can be formulated. Firstly, it can be predicted that participants in this study will be more accurate in detecting words which have patterns of phonotactics found in both English and Russian, and this ability will increase over sessions. Secondly, it can be predicted that participants in this study will perform better in detecting words which are stressed on the first syllable, and this ability will also increase over sessions.

All predictions in this thesis are summarised at the end of Chapter 5, which is dedicated to a discussion of the methodology used in the present study.

Chapter 5. Methodology

5.1 Participants

Twenty-nine students from Newcastle University, UK were recruited using an advertisement. The advertisement (see Appendix B2) was circulated among Newcastle University students: (1) by email with the help of secretaries from different schools; (2) hard copies were hung on corkboards in various schools as well as in Newcastle University Student Union; and (3) a researcher distributed the printed mini brochures of the advisement to passers-by. The advertisement specified that the researcher was looking for native speakers of English without knowledge of Slavic languages to take part in a *linguistic* experiment, the aim of which was to investigate how foreign languages are learned. To be eligible to participate in the experiment, participants had to identify themselves as having no knowledge of Slavic languages. The advertisement also mentioned that participants would be reciprocated with a £10 Amazon voucher, and there would be biscuits, sweets, and chocolates throughout the experiments. Moreover, the advertisement clarified what would be expected from participants, i.e. that on each day participants would be required to listen to an audio file of an unknown language and then do listening tasks on a laptop and that on a final day they also would need to complete a short bibliographical questionnaire.

All twenty-nine participants reported no known hearing or language impairment[43]. All participants were native speakers of English, but one participant was English-Welsh bilingual from birth[44]. The mean age of *all* participants was 23 years and three months (*SD*=6.50 months), the minimum age was 18 years old, and the maximum age was 43 years old. One participant out of twenty-eight did not provide his age. Twenty-one participants were female, and seven were male. Twenty participants did an Undergraduate degree, four participants did an MA degree, and four

[43] Two participants reported that they had a learning difficulty, but their results were not excluded. One participant (coded as part27 in answers on questionnaire, see Appendix 8.14) was deleted from the data analysis as she did a BA honours degree in French and Russian in 1991 although she reported that she forgot most of it as she has never used it after graduation.
[44] Background of these participants, as well as more details about these students is in Appendix 8.14

participants were on the first stage of their PhD degrees, all at Newcastle University. Cantonese, Mandarin Chinese, French, German, Japanese, Korean, Spanish and Portuguese were reported as L2s and in a few cases L3s which participants knew. All participants specified that they started learning these languages after the age of 10 through instruction, but two participants began learning L2s at the age of 7. For one participant (part8) it was Spanish which she learnt naturalistically as she resided in Spain from the age 7 to 16, and for another one (part12) it was French which she studied through instruction. Both participants were undergraduate students of linguistics. Additionally, two more participants (part17), and (part20) reported that they learned naturalistically Italian and Cantonese. The first of these learned some limited Italian for travelling, and the second one learned some Cantonese while living in Hong Kong. The average score of participants on their L2 and L3 language(s) ability for speaking, listening, writing, reading, grammar and pronunciation skills was 2.63 (*SD*=1.12), the lowest score was 1, and the highest score was 5. These scores were self-rated on a scale from 1 to 5 where 1 means *not good,* and 5 means *very good*.

One of the aims of this study was to investigate if there would be a difference between linguistically naïve and linguistically sophisticated groups in their ability to respond to new words after minimal input. Therefore, to answer this research question, all participants were broken into two linguistic groups which had two levels (i.e. *linguistically sophisticated group* vs *linguistically naïve group*). The linguistically sophisticated group contained 15 students which comprised of students of either a degree with a linguistics component or a degree with a language component. Additionally, one participant, although not from a linguistic degree, was assigned to this group because in the open question of the bibliographic questionnaire she indicated that her father was a professor of linguistics who can speak Spanish, specifying that she knew Spanish as a foreign language and rated her L2 skills overall 4.33 out 5. A score about this high was found in two participants (part 8 and part 11) one of whom did a degree in linguistics and the other a degree with a foreign language component. The mean age of the linguistically sophisticated group was

22.93 (*SD*=4.78) with a minimum age 18 and a maximum age of 34, all participants but one were female. The linguistically naïve group contained 13 students who studied non-linguistic and non-language related degrees. The mean age of the linguistically naïve group was 23.83 (*SD*=8.29) with a minimum age 18 and a maximum age 43, where seven participants were female and six were male. Lastly, the list of participants and their characteristics as collected from the bibliographic questionnaire is included in Appendix 8.14.

5.2 Materials

The main aim of the study was to test whether or not an adult, without formal instruction, can segment the speech stream of a completely unfamiliar natural language that was presented in the form of an aural input for seven minutes on four consecutive days. It should be recalled that the present study in particular aimed to investigate the effects of the following on speech segmentation: (i) effects of single cues (phonotactics, stress and word length); (ii) effects of combinations of cues (i.e. phonotactics with stress, and phonotactics with word length); (iii) effect of learning over time; (iv) effect of single cues over time; (v) effect of generalisation; also (vi) how all of these effects just listed are realised in linguistically sophisticated as opposed to linguistically naïve participants who were tested in a word recognition task and a forced-choice task. Furthermore, the present study aimed to control whether participants were paying attention to the input by implementing a cognate-identification task, which also tested the recognition of cognates on all four days (all tasks and their purpouse are described in Section 5.3). The present study was designed specifically to address these aims. Section 5.2.1 focuses on describing stimuli of the present experiment by describing targets and distractors, and Section 5.2.2 will review how input was recorded and how cognates were selected and used to record additional sentences.

5.2.1 Stimuli

All aural stimuli of the present experiment were recorded by a female native speaker of Standard Modern Russian in a soundproof booth using an Edirol R-44 recorder with default level settings of

16 bit 48 KHz. Each word was recorded individually with neutral intonation. The word-initial and final silence was removed using PRAAT version 6.0.28. Digitised versions of the stimuli were saved on the researcher's laptop for playback during the experiment.

All stimuli were selected in such a manner that monosyllabic words were of CCVC structure, and bisyllabic words were of CCV.CVC structure, but few times words of CCVC.CVC structure were used. This is because it was aimed to have all words that were real Russian words, but it was not possible to have all words which satisfied both CCV.CVC structure and real word status, therefore other real words of Russian were used which were of CCVC.CVC structure. All these words were taken from the electronic dictionary *multiran.ru,* which has corpus properties by allowing searches for words according to specific criteria. Stimuli were never minimal pairs, as minimal pairs might be harder to recognise (Carroll & Windsor 2015). An effort was made for the stimuli to be nouns because when targets were embedded into a text, it was easy to create a semantically well-formed sentence. Moreover, the influence of the phonetic inventory was not controlled, which means that all stimuli and experimental sentences contained sounds which are specific to the Russian language (not found in English). Throughout the process, all stimuli were checked by a native speaker of English, so they did not resemble existing English words.

5.2.1.1 Targets

To see how participants would respond to words, phonotactics of which can be transferred from their L1 as opposed to how participants would react to the novel words with phonotactics of Russian, two lists of onset consonant clusters (CC type) were selected which were called experimental condition *phonotactics* with levels *native* vs *non-native*. The former list was required to test the effect of pre-existing knowledge of the L1 (English) and the latter was needed to examine how participants would respond to new phonotactics in L2 (Russian).

The native list (*n*=24) had eight types of consonant clusters which frequently occur in both English and Russian languages. All these phonotactics had a rising sonority and were obstruents followed by liquids (MSD=2):

kl-, bl-, gl-, sm-, sl-, pl-, kr-, gr-.

The non-native list (*n*=24) were chosen on the assumption that they would require the learners to create novel sound forms which are existent only in Russian. The list also contained eight consonant clusters which are frequent in Russian but do not occur in English. Some of these phonotactic types had a rising sonority (MSD=2), and others had sonority plateaus (MSD=0), and none of these phonotactic types violated Russian:

hl-, kn-, sv-, ʃt-, tv-, ʃk-, zv-, sr-.

These two lists were used as a basis for constructing targets (words with these consonant clusters) which participants heard in the input. Next, these targets were further subdivided into *monosyllables* and *bisyllables* of a variable *length* of a word. It was an experimental condition to investigate what effect length cue has on a participants' ability to respond to new words. As a result, there were eight monosyllabic targets and 16 bisyllabic targets in each native and non-native phonotactic condition. Furthermore, in order to test if participants were affected by native prosody, mainly relying on MSS to segment words of Russian, the 16 bisyllabic words were further divided into words which were stressed on the first syllable (*strong-weak*) and words which were stressed on the second syllable (*weak-strong*) to reflect an experimental condition *stress*. Crucially, recall that the present study did not attempt to investigate whether participants would show sensitivity to a particular stress placement in Russian due to its complex stress assignment, as this was discussed in detail in Chapter 4. However, due to the 'free' stress placement within a stem of Russian polysyllabic words, it was possible to come out with two lists of real Russian bisyllabic words, the first of which was comprised of words stressed on the first syllable, the second of which of words stressed on the second syllable. They were eight items with strong-weak stress, and eight items with weak-strong stress in each phonotactic condition. Considering all experimental conditions of the present study (nativeness, word length, and stress), there were minimally eight items and maximally 24 items per experimental condition which generated 48 targets overall which are illustrated in Table 5-1. Eight items per condition were selected because it is generally considered to be the minimum number of trials necessary to obtain a reliable statistically power; although there is no consensus among researchers on how many items/trials should be included in the experiment (Boudewyn *et al.* 2017).

	Phonotactic condition = Native							
	Monosyllabic words			Bisyllabic words				
				Strong-weak words			Weak-strong words	
	IPA	Russian		IPA	Russian		IPA	Russian
1	[klʲik]	klik	2	[ˈklʲevʲɪr]	klever	3	[klaˈtʃʲ ok]	klochok
4	[blʲef]	blef	5	[ˈblʲinʲɪk]	blinnik	6	[blʲɪzʲˈnʲets]	bliznec
7	[glas]	glaz	8	[ˈglʲænʲɪts]	glianec	9	[glaˈtok]	glotok
10	[smʲesʲ]	smes’	11	[ˈsmʲenʲʃʲːɪk]	smeschik	12	[smʊˈglʲak]	smugliak
13	[slux]	sluh	14	[ˈs⁽ʲ⁾lʲitək]	slitok	15	[slaˈvarʲ]	slovar’
16	[plof]	plov	17	[plʲedʲɪk]	pledik	18	[plaˈtok]	platok
19	[krax]	krah	20	[ˈkrolʲɪk]	krolik	21	[kraˈvatʲ]	krovat’
22	[grom]	grom	23	[ˈgruʃʲːɪk]	gruzschik	24	[graˈfʲin]	grafin
	Phonotactic condition = Non-native							
	Monosyllabic words			Bisyllabic words				
				Strong-weak words			Weak-strong words	
	IPA	Russian		IPA	Russian		IPA	Russian
25	[xlʲep]	hleb	33	[ˈxlopʲɪts]	hlopec	41	[xlaˈpok]	hlopok
26	[knʲelʲ]	knel	34	[ˈknʲigəm]	knigam	42	[knʲaˈzʲok]	kniazek
27	[svʲet]	svet	35	[ˈsvʲitək]	svitok	43	[svʲɪˈnʲets]	svinec
28	[ʃtat]	shtat	36	[ˈʃtopər]	shtopor	44	[ʃtʊrˈval]	shtyrval
29	[tvʲɪt]	tvid	37	[ˈtvorək]	tvorog	45	[tvaˈrʲets]	tvorec
30	[ʃkaf]	shkaf	38	[ˈʃkolʲnʲɪk]	shkolnik	46	[ʃkodˈnʲɪk]	shkodnik
31	[zvuk]	zvyk	39	[ˈzvonaˈrʲ]	zvonar’	47	[zvaˈnok]	zvonok
32	[srok]	srok	40	[ˈsrubʃʲːɪk]	srubschik	48	[srasˈtok]	srostok

Table 5-1. Targets.

5.2.1.2 Distractors (generalisable new items)

Just as the total number of targets, the first set of distractors were *generalisable* items, which were 48 items. They were all words with the same phonotactics as the targets, but unlike the targets, they were new items, as they did not occur in the input. They were selected and partitioned concerning experimental conditions in the same way as the targets, respecting the numbers of items. That is there were 24 items in each phonotactics condition, and each phonotactic condition contained eight monosyllables, and 16 bisyllables, with bisyllabic words, further broken into 8 with strong-weak stress pattern and eight with a weak-strong stress pattern. The intention behind this group of distractors (generalisable new items) was to see whether participants had picked up phonotactic properties from the input and were able to generalise these to new items. It was anticipated that accuracy on these generalisable new items would be similar to that of the targets. These distractors are illustrated in the following Table 5-2.

	Phonotactic condition = Native							
	Monosyllabic words			Bisyllabic words				
				Strong-weak words			Weak-strong words	
	IPA	Russian		IPA	Russian		IPA	Russian
49	[klat]	klad	57	[ˈklapən]	klapan	65	[klʊˈbok]	klubok
50	[blat]	blat	58	[ˈbludnik]	bludnik	66	[blaˈtʃʲok]	blachok
51	[glupʲ]	glub’	59	[ˈglazʲɪk]	glazik	67	[glaˈgol]	glagol
52	[smok]	smog	60	[ˈsmʲe ʒʲ:nʲɪk]	smezhnik	68	[smarˈtʃʲok]	smorchok
53	[s⁽ʲ⁾lʲisʲ]	sliz’	61	[ˈslonʲɪk]	slonik	69	[s⁽ʲ⁾lʲɪzʲˈnʲak]	slizniak
54	[plaʃʲ]	plasch	62	[ˈplotʲɪk]	plotik	70	[plaˈmbʲɪr]	plombir
55	[krʲik]	krik	63	[ˈkrovnʲɪk]	krovnik	71	[krʲɪˈkun]	krikyn
56	[grʲasʲ]	griaz’	64	[ˈgroxət]	grohot	72	[grɨˈzun]	gryzyn
	Phonotactic condition = Non-native							
	Monosyllabic words			Bisyllabic words				
				Strong-weak words			Weak-strong words	
	IPA	Russian		IPA	Russian		IPA	Russian
73	[xlor]	hlor	81	[ˈxlʲupʲɪk]	hlupik	89	[xlʲɪˈvok]	hlevok
74	[knʲot]	knet	82	[ˈknutʲɪk]	knutik	90	[knʲi ʒʲ:nʲɪk]	knizhnik
75	[svot]	svod	83	[ˈsvoraˈt]	svorot	91	[svʲɪˈstok]	svistok
76	[ʃtuk]	shtyk	84	[ˈʃturmən]	shturman	92	[ʃtɨˈrʲok]	shtyrek
77	[tvʲil]	tvil	85	[ˈtvʲistər]	tvistor	93	[tvʲorˈdos]	tverdoz
78	[ʃkʲif]	shkiv	86	[ˈʃkʲipʲɪr]	shkiper	94	[ʃkaˈlʲar]	shkoliar
79	[zvʲerʲ]	zver’	87	[ˈzvʲozdəm]	zvezdam	95	[zvaˈnʲets]	zvonec
80	[srɨf]	sryv	88	[ˈsrʲestʃʲɪk]	srezchik	96	[sramˈnʲɪk]	sramnik

Table 5-2. Generalisable distractors.

5.2.1.3 Distractors (non-generalisable new items)

The second set of distractors were new *non-generalisable* items which contained an additional 48 items, which included onset phonotactics utterly different to what participants encountered during the input. Given that the onset phonotactic property of these non-generalisable distractors did *not* appear in the input, it was anticipated that the accuracy (i.e. participants ability to reject these words as non-targets) would be higher on the non-generalisable items than the generalisable ones. Similarly, to the generalisable distractors, this group of distractors needed to satisfy *phonotactic* experimental conditions of the present study by having 24 items with native phonotactics; and 24 items with *non-native* phonotactics. To satisfy the native phonotactics condition, 24 real words of Russian were selected. All these words contained phonotactics highly frequent in both English and Russian languages (in fact common cross-linguistically), but once again, none of these appeared in the input phase. All these phonotactics had a rising sonority (MSD=2 and MSD=1):

sk-, tr-, br-, dr-, fl-, fr-, sp-, sn-.

However, it was much harder to satisfy the non-native phonotactic condition because the researcher needed to come out with another eight instances of CC clusters which were supposed to start with CC sequences with different phonotactics to what participants heard during the input but still needed to be legal in Russian. The selection of real Russian words which would satisfy the criteria of this experimental condition proved to be possible only to the extent of identifying four clusters of Russian. They were highly infrequent Russian words, where one word had a sonority plateau (MSD=0), and the other three violated sonority (MSD=-1, and MSD=-2):

rt-, lʒ-, lg-, ptʃ-.

The other four clusters were selected by thinking outside of what is allowable in Russian and English but possible universally. All these clusters had a sonority plateau (MSD=0). These clusters are illustrated below:

mp-, gb-, nk-, ht-.

So, eight clusters in total were used to come out with the list of 24 items which were nonsense words. As this list of words contained clusters which were illegal in Russian, it was decided to name this group of clusters as *neither* to reflect the idea that these clusters did not appear in the input and did not conform to the phonotactic expectation of what is possible in English and Russian. They were partitioned into three lists using the same criteria of experimental conditions such as length and stress just as was done for the targets and generalisable distractors. That is there were 8 monosyllables, 16 bisyllables. The bisyllables were further broken into 8 with a strong-weak stress pattern and 8 with a weak-strong stress pattern. These non-generalisable distractors are shown in the following Table 5-3.

	Phonotactic condition = Native							
	Monosyllabic words			Bisyllables				
				Strong-weak words			Weak-strong words	
	IPA	Russian		IPA	Russian		IPA	Russian
97	[skas]	skaz	105	[ˈskupʃʲːik]	skupschik	113	[skaˈkun]	skakun
98	[trʲelʲ]	trel’	106	[ˈtrʲepʲɪt]	trepet	114	[ˈtravnʲɪk]	travnik
99	[brak]	brak	107	[ˈbratʲɪk]	bratik	115	[brʊˈsok]	brysok
100	[dropʲ]	drob’	108	[ˈdrotʲɪk]	drotik	116	[drʊˈʐok]	druzhok
101	[flʲus]	flus	109	[ˈflotʲɪk]	flotik	117	[frʲɪˈgat]	fregat
102	[spʲex]	speh	110	[ˈsposəp]	spasob	118	[sparˈnʲik]	sparnik
103	[sʲnʲek]	sneg	111	[ˈsʲnʲimək]	snimok	119	[snaˈʃʲːik]	snoschik
104	[prut]	prud	112	[ˈprʲibɨlʲ]	probyl’	120	[praˈʃɨf]	proshiv
	Phonotactic condition = None							
	Monosyllabic words			Bisyllables				
				Strong-weak words			Weak-strong words	
	IPA	Russian		IPA	Russian		IPA	Russian
121	[mpar]	mpar	129	[ˈmpovər]	mpovar	137	[mpaˈrʲik]	mparik
122	[gbʲit]	gbit	130	[ˈgbagʲet]	gbager	138	[gbʲɪˈnom]	gbinom
123	[nkʲib]	nkib	131	[ˈnkomak]	nkomak	139	[nkaˈmʲin]	nkamin
124	[xtʲex]	hteh	132	[ˈxtʲerʲɪk]	hterik	140	[xtaˈnok]	htonok
125	[rtutʲ]	rtut’	133	[ˈrtovʊn]	rtovun	141	[rtʲɪˈʃʲːɪˈk]	rtischek
126	[lʒets]	lzhec	134	[ˈlʒɨvən]	lzhivon	142	[lʒeˈmud]	lzhemud
127	[lgatʲ]	lgat’	135	[ˈlgunʲam]	lguniam	143	[lgaˈnʲiʃ]	lganish
128	[ptʃak]	pchak	136	[ˈptʃɪvʲer]	pchiver	144	[ptʃʲɪˈlʲak]	pcheliak

Table 5-3. Non-generalisable distractors.

5.2.2 Input and cognates

Two lists of different sentences were constructed around all 48 targets. Each list contained 48 different sentences, with each target being used only once. These sentences with targets, just as with the targets themselves, were not related in meaning (but see the motivation behind using two lists later). The sentences contained a minimum of four and a maximum of six words. Previous research, e.g. Shoemaker & Rast (2013) found the initial and final words of an utterance were recognised better than those in a medial position as a result of sensitivity to the edges of prosodic constituents. Therefore, to attribute the segmentation effect in this study to the influence of the variables which were put under investigation, namely *phonotactics*, *stress* and *length* each target appeared only in the medial sentence position, and it never occurred anywhere else in a list of 48 sentences. What was meant by the medial sentence position is that a target was not the first word or the last word in a sentence, and it was roughly followed and proceeded by three to five syllables. Given that Russian masculine nouns do not change for gender and case in the nominative and accusative cases, and the flexible word order of Russian helped to position the targets in the medial sentence position. Gomez (2002) in his study on an artificial language has shown that even short pauses can cue a word boundary, so in the present study, care was taken to pronounce each word within a sentence to eliminate any pauses, and also in such a manner that none of the words received a focal accent. Additionally, to make sure that the effect of consonant clusters was definitely due to the onset clusters of target words, and not the cumulative effect of encounters with these clusters throughout the input, a care was taken so all other words in the input which were not targets contained no clusters of target words. Moreover, each target appeared after a word ending with phonemes [n], [m], [l], [t], [oj] or [ij], e.g. *zolotoj slitok* [zəlatˈoj slʲˈitək] ‘gold bar’. Hence, in conjunction with the preceding segment, the first segment of a target always signalled a word onset boundary in both English and Russian, for instance, for both languages:

**t#kl-, *t#bl-, *t#gl-, *t#pl-,*j#sm, *j#sl, *j#kr-, *j#gr-*

Similarly, to Gullberg *et all*. (2010, 2012) the text was recorded to respect the properties of coherent discourse as much as possible. Additionally, care was taken that all other content words in sentences (not-target words) did not occur too frequently, but no special method was used to check the extent to which it was true. Each list of sentences was randomised using the Excel randomisation option.

The reason why two lists of 48 sentences (but not one list) were created was because it was planned to intersperse each list of 48 sentences with sentences containing English – Russian cognates, where 11 sentences containing cognates relevant to the theme "music" were inserted into the first list of 48 sentences, and the second list containing 11 cognates relevant to the theme "university life" were added to the next list of 48 sentences. Inserting sentences with cognates was done to be able to check if participants were paying attention while listening to the input. According to Milroy (1909) fatigue can affect the dependent variable, and fatigue was particularly an issue in the present study because of a within-subject design. It was anticipated that sentences with cognates would activate L1 lexical words, and therefore would make participants more interested in listening to the input, due to them perceiving it as if they were getting the gist of what the input was about. Whether participants were paying attention to the input was tested by seeing if the recognition of cognates would improve with increasing input in a cognate identification task which are described in Sections 5.3.1.3 and 5.3.3.1. Provided the robustness of the effect of cognates as discussed in Chapter 3, it was expected that if participants do not perform well on the cognate identification tasks, it is likely because they were not paying sufficient attention to input. The researcher relied on her intuitions when selecting Russian-English true cognates, which were used for targets and distractors.

As already stated above, two lists of true English-Russian cognates were created. See Appendix A5. for the lists of cognates which were used in the cognate identification task. The first list included words which matched the theme 'music', for example [ˈlʲirʲɪkə] "lyrics" and [bərʲɪˈton] “baritone”; and the second list contained words which matched the theme ‘university life’, for example [stuˈdʲent] “student” and [unʲɪvʲɪrsʲɪˈtʲet] “university”. 11 sentences were constructed with each set of cognates. These sentences were used to break the sequence of two lists of randomised sentences of the main input, which contained targets. A sentence with a cognate was inserted after every four sentences so that the first list included all the sentences about *music* and the other list comprised all the sentences about *university life*. Participants listened to the list containing cognates about music (n=48 sentences with targets + 11 sentences with cognates about “music”) on the first and second day; and they listened to another list containing cognates about “university” (n=48 sentences with targets + 11 sentences with cognates about university) on the third and fourth days. It means that participants listened to one input on the first and second day, and different input on the third and fourth day, but exposure to targets words (see Section 5.2.1.1) was consistent across the sessions.

Both lists of input were audio-recorded by the same female speaker of Russian in a sound-proof booth using Edirol R-44 recorder with the default level settings of 16 bit 48 KHz. The sentences were recorded with a normal intonation, with the gap between sentences of five seconds. The duration of the passages were 3 minutes 31 seconds for the one containing cognates about “music” (see Appendix A.1.1), and 3 minutes 40 seconds for the one containing cognates about “university” (see Appendix A.1.2). Digitised versions of the recordings were saved on the researcher’s laptop for playback during the experiment.

5.3 Procedures and tasks

All participants were required to meet with the researcher in a quiet room in the Percy Building in Newcastle University except one participant[45]. Before coming to the first session, participants emailed the researcher to express an interest. Every participant was asked to book four sessions for

[45] One participant was tested in a Phonetic lab (room 2.13) in King George VI Building, Newcastle University as it was arranged specifically just for her.

30 minutes on consecutive days. Participants were asked to come on four successive days, which was done to explore the effect of the words' detection over time. Good quality equipment such as the researchers' *Dell* laptop (model Inspiron 13 5000) and comfortable headphones *Sony* (model MDR-ZX110AP) with a self-adjusting headband and cushioned ear-pads was used. Each participant was tested individually. If participants completed four days of an experiment, they received an *Amazon thank-you* card with a voucher value worth £10. Throughout each experiment, participants were offered biscuits and chocolates. It was made sure that on each day, participants felt comfortable with the equipment, anybody could adjust the headband of headphones and the sound volume to their preferences. A week or two after an experiment finished, each participant was sent graphs with their accuracy results on all four days and a quick explanation of their achievements, everyone was invited to get in touch with a researcher if they wanted to discuss their results.

The present study utilised a *word recognition task*, a *forced-choice task* which was always the same on all sessions; and different versions of the *cognate identification task*, the first one utilising cognates about "music" used on the first and second day, and the second version with cognates about "university life" was used on the third and fourth day. All experimental tasks were created using experimental software *OpenSesame* version 3.1.6 Jazzy James (Mathôt *et al.* 2012). It is a free and open-source programme specifically designed to create experiments for psychology and neuroscience. The programme benefits from the user-friendly graphical interface which involves dragging distinct experimental units from the item toolbar and dropping them onto the overview timeline area to create an experiment, which is straightforward even for those without programming experience. All participants took part in all experimental tasks on each of four days.

The next section is structured to review the procedure of the experiment in a chronological manner, with the details about the word recognition task, the forced-choice task. The first version of the cognate identification task which tested cognates about 'music' are discussed in session 1

(Section 5.3.1), the second version of the cognate identification task which tested cognates about 'university life' is summarised within session 3 (Section 5.3.3). The hypotheses and predictions of this study are presented in Section 5.4.

5.3.1 Session 1

On day 1, participants were asked to sign an ethical consent form which outlined the procedure of the experiment and information about the confidentiality of the obtained data, and the experimental schedule was confirmed. It was not made clear until the end of the experiment which foreign language participants were going to be exposed to except that they knew it was a natural language and it was one of the Slavic languages, and that they would find out which language it was when they completed all days of the study. The researcher explained that even if participants had some guesses about what language it was, they were asked not to check any facts or get additional exposure to any Slavic languages while they were taking part in the study. After that, participants were instructed to listen to and try to make sense of the first audio-file (with the theme 'music') that was played to them twice (see Appendix A.1.1). It was decided to play audio recording twice (which made a total of seven minutes, two seconds of exposure) as it appeared to be sufficient for showing effects of words' recognition after seven minutes of Mandarin input (Gullberg *et al.* 2010, 2012).

5.3.1.1 Word Recognition Task

After the listening phase, participants took the word recognition task, which had two purposes. The first purpose of the word recognition task was to see if participants could detect words (targets) they had heard in the input phase as opposed to words (distractors) they had not heard and to examine which cues participants relied on for word detection over four consecutive days. The second purpose of the task was to investigate if participants could extend the knowledge of phonotactic patterns they had heard in the input to the new items they had not heard in the input, that is by comparing their accuracy on *targets* as opposed to novel words with *generalisable* and *non-generalisable* properties. It was expected, if *ab-initio* learners have generalisation capacities, accuracy on generalisable distractors would be similar to targets, and accuracy on non-generalisable distractors will be the highest among three levels of the type of stimuli condition because participants will be rejecting non-generalisable distractors accurately.

Similarly, to the procedure described in Gullberg *et al.* (2012) participants sat approximately 60 cm from a computer screen and the experimental list of audio files was played *via* headphones, one at a time. The experiment started with the presentation of the sketchpad on the screen, which showed the following instructions:

"You will listen to 144 words presented to you one by one.
If you think:
-you heard the word previously, press 'z';
-you did not hear the word previously, press 'm'.
Be as fast and accurate as possible (If you do not respond for 4 secs, the next item will be automatically played).
Press any button to begin".

The white fixation dot appeared at the centre of the screen precisely for the duration of each sound file. This was followed by presenting the text on the screen, together with the number of trials (1-144), a green letter 'z' on the left-hand side of the screen, and a red letter 'm' on the right-hand side of the screen for up to 4 seconds (the maximum allowed time for participants to respond by pressing a button). Upon a keypress, participants heard a beeping noise. To prevent this noise from overlapping with the presentation of the next stimuli, a black fixation dot appeared at the centre of the screen for 0.5 seconds to signal the end of the trial. The presentation of the sound files had been randomised along with the variables they were associated with just once, and the sound files played sequentially by the software identically for each participant on all four days. The randomised version of the experimental design for the word recognition task is in Appendix A.2. When the final item played, the text appeared on the screen which read that it was the end of the block, and participants were thanked for participating and were asked to press any key to exit the program. The logger tool was used to record all the variables for each participant.

The word recognition task contained all 144 stimuli, 48 of which were target items (required the ‘z’ response); and 96 of which were distractor items (required the ‘m’ response). It took no longer than eight minutes to complete this task.

5.3.1.2 Forced-Choice Task

The forced-choice task was taken straight after the word recognition task. The primary motivation for running an additional task (the forced-choice task) was because it was anticipated that the results from the word recognition task might show only limited sensitivity because the word recognition task could have been too difficult due to 144 stimuli presented to participants one by one requiring participants to respond according to “feel”. Whereas it was expected that the forced-choice task would have been easier as the nature of the task involves comparing only 48 pairs of stimuli, one of which is a target, and another one is a distractor. Consequently, participants knew that one of the words ought to be a target.

The purpose of the forced-choice task was to follow up on findings from the word recognition task, that is to see if participants could detect targets from words that they did not hear in the input, and to examine which cues participants relied on for word detection. As for the distractor items, there was a choice whether to utilise the list of generalisable distractors or the list of non-generalisable distractors. It was decided to use the generalisable distractors because a choice between two stimuli (where one is a target, and another one is a distractor with a generalisable property) would allow testing if learners could generalise to phonotactic properties heard in the input. That is, if *ab-initio* learners have generalisation capacities, participants would not distinguish between targets and generalisable distractors on the forced-choice task.

The forced-choice task was also created in OpenSesame and followed the same general procedure as the word recognition task. The task started with the presentation of the following instructions on the screen:

"You will listen to 48 pairs of words presented to you one by one.

If you think:

- you heard the first word in the recording, press 'z';

- you heard the second word in the recording, press 'm'.

Be as fast and accurate as possible (If you do not respond for 4 secs, the next item will be automatically played).

Press any button to begin".

The participants were required to listen to 48 pairs of items presented one by one. The task was to decide which of two items in each pair they had heard in the input. If participants thought they had heard the first word, they were asked to press 'z', and to press 'm' if they thought it was the second word. Participants were asked to respond as fast as possible, and if they failed to respond within four seconds, the program moved to the next item.

Each pair of words was presented as follows. The white fixation dot appeared at the centre of the screen for 500 milliseconds. After which the text 'WORD 1' was displayed at the centre of the screen and a sound file simultaneously played, which was followed by the text 'WORD 2', and the other sound file was simultaneously played. The sequence of two sound files was delimited by presenting a blank screen for 0.5 seconds. Then the sketchpad appeared showing the number of the trial in the centre of the screen and the two words, on the left- and right-hand sides of the screen respectively, and the participants were given four seconds to make a choice.

48 pairs of sound files consisted of the list of targets from the input on the one hand, and the list of generalisable distractors on the other hand, which were matched by experimental conditions (variables): phonotactics, words' length and syllable stress. For instance, the distractor counterparts to the targets

klik [klɨk], *klever* [ˈklʲevʲɪr] and *klachok* [klaˈtʃʲ ok]

were

klad [klat], *klapan* [ˈklapən] and *klubok* [klʊˈbok].

The display of stimuli pairs was randomised, and the sound files played sequentially by the software identically for each participant on all four days. The randomised version of the experimental design for the forced-choice task is in Appendix A.3. When the final combination of sound files played, a message that it was the end of the block appeared on the screen, and the participants were thanked for participation and asked to press any key to exit the programme. The logger tool was used to record all the variables for each participant. This task took no longer than 5 minutes to complete.

5.3.1.3 Cognate Identification Task (about music)

Finally, in order to test if participants were paying sufficient attention to the input by detecting cognates that they heard in the input as opposed to other cognates they did not hear, they took the *first* version of the *cognate identification task* because on the first and second day participants listened to input which was interspersed with sentences containing cognates belonging to the theme "music". The cognate identification task was also created in OpenSesame. The task started with the presentation of instructions on the screen. The instructions to the cognate identification task were the same as those of the word recognition task. That is, participants read that they would listen to 20 words presented to them one by one and if they thought they had heard the word before, they needed to press 'z' or otherwise to press 'm' on keyboard. As with both tasks above, participants were given 4 seconds to respond. The test items included 10 cognate words which were from the audio recording, and 10 distracters which were other cognates between English and Russian but were semantically unrelated to each other and did not appear in the treatment recording, for example, *bariton* [bərʲɪtˈon] 'baritone' and *pianino* [pʲɪanʲˈinə] 'piano' were the targets, and *budzhet* [bʲudʒˈɛt] 'budjet' and *hokhej* [xakʲˈej] 'hockey' were distracters. All testing items were randomised, and the sound files played sequentially by the software identically for each participant on all four days. The randomised version of the experimental design for the cognate identification task is in Appendix A.4. This task took no longer than two minutes to complete. The whole session on the day one lasted about 30 minutes.

5.3.2 Session 2

The treatment and experimental tasks on the second day were the same as those on the first day. The whole session lasted about 30 minutes.

5.3.3 Session 3

On the third day participants were instructed to listen to and try to make sense of the second audio-file (with a theme ‘university life’) that was played to them twice, which made a total of seven minutes 20 seconds of exposure) The whole session lasted about 30 minutes. The rest of the session on day three is described in the next section.

5.3.3.1 Cognate Identification Task (about university life)

After participants listened to the input, they took the word recognition and the forced-choice tasks, which were precisely the same as on previous days. Also, participants took the *second* version of the *cognate identification task*, which was identical in its design to that of the first and second day (see Section 5.3.1.3) for the description of what the design involved, and Appendix A4.), but it had different targets and distracters. The targets were 10 cognate words from the audio recording ‘university life’, for instance *institut* [ɪnʲsʲtʲɪtˈut] ‘institute’ and *student* [stʊdʲˈent] ‘student’, whereas distractors were the other 10 cognates which were semantically unrelated and did not appear in the input, such as *futbol* [fʊdbˈol] ‘football’ and *komjuter* [kampʲjˈutɨr] ‘computer’.

5.3.4 Session 4

The treatment and experimental tasks on the fourth were the same as those on the third day. Participants also completed *language history questionnaire* (see Appendix B.1) adapted from

(Gullberg & Indefrey 2003) to collect bibliographical data. The most important information which was collected from these questionnaires was already incorporated into the description of participants of the present study (in Section 5.1). Answers on language questionnaire are provided in Appendix B.4. The whole session on the fourth day lasted about 30 minutes.

5.4 Research questions and hypotheses

To reflect the main aim of the present study of how *ab-initio* second language learners start to detect words in Russian, the research questions and hypotheses of the present study can be divided into five main categories. In particular, (1) exploring the effect of learning over time; (2) exploring the effect of single cues (phonotactics, stress, and word length); (3) exploring the effect of combinations of cues (phonotactics and stress, and phonotactics and word length); (4) exploring the effect of each cue (phonotactics, stress, and word length) over time; (5) exploring the effect of generalisation; and finally (6) how all of these effects just mentioned are realised in linguistically sophisticated participants as opposed to linguistically naïve participants. Additionally, each research question (RQ) below, along with its related hypothesis, is followed by a more specific prediction which has already been discussed throughout the thesis.

(i) Exploring effect of learning over time (effect of session)

1. Does learners' ability to detect Russian words from the input increase over sessions?

Hypothesis 1: There will be differences in participants' accuracy between sessions.

In particular, participants' accuracy on targets will improve with an increased amount of input.

(ii) Exploring effect of single cues (phonotactics, stress, word-length)

2. Do learners rely on L1 phonotactics, or they develop sensitivity to Russian phonotactics when detecting words from the input?

Hypothesis 2: There will be differences in participants' accuracy among words with *native*, *non-native*, and *neither* phonotactics.

In particular, participants of this study will be more accurate in detecting targets which have native phonotactics than non-native phonotactics.

3. Do learners rely on MSS (strong-weak stress pattern), or they rely on weak-strong stress pattern when detecting novel words from the input?

Hypothesis 3: There will be differences in participants' accuracy between *strong-weak* words and *weak-strong* words.

In particular, participants of this study will be more accurate in detecting targets which are stressed on the first syllable, than on the second syllable.

4. Do learners show preference to bisyllabic over monosyllabic words when detecting words from the input?

Hypothesis 4: There will be differences in participants' accuracy between *bisyllabic* and *monosyllabic* words.

In particular, participants of this study will be more accurate in detecting bisyllabic targets than monosyllabic targets.

(iii) Exploring effect of combinations of cues

5. Are learners guided by an interaction between phonotactics and MSS when detecting words from the input?

Hypothesis 5: There will be differences in participants' accuracy between the interaction of *phonotactics* and all *stress* of words.

In particular, participants of this study will be more accurate in detecting targets with native phonotactics, and which are stressed on the first syllable, than detecting targets with native phonotactics and word-final stress.

6. Are learners guided by an interaction between phonotactics and word length when detecting words from the input?

Hypothesis 6: There will be differences in participants' accuracy between the interaction of *phonotactics* and *length*.

In particular, participants of this study will be more accurate in detecting targets which have native phonotactics than detecting words with non-native phonotactics; and there should be a preference for bisyllabic targets over monosyllabic targets.

(iv) Exploring effect of each cue (phonotactics, stress, length) over time

7. Does sensitivity to phonotactic constraints in the detection of words from the input increase over sessions?

Hypothesis 7: There will be differences in participants' accuracy between the interaction of *phonotactics* and *session* conditions.

In particular, participants of this study will be more accurate in detecting targets which have native phonotactics than detecting targets with non-native phonotactics, and this ability will increase over sessions.

8. Does sensitivity to MSS (strong-weak stress pattern) in the detection of words from the input increase over sessions?

Hypothesis 8: There will be differences in participants' accuracy between the interaction of *stress* and *session* conditions.

In particular, participants of this study will be more accurate in detecting targets which are stressed on the first syllable than on the second syllable, and this ability will increase over sessions.

9. Does sensitivity to word length in the detection of words from the input increase over sessions?

Hypothesis 9: There will be differences in participants' accuracy between the interaction of *length* condition and all *sessions*.

In particular, participants of this study will be more accurate in detecting bisyllabic targets than monosyllabic targets, and this ability will increase over sessions.

(v) Generalisation

10. Can learners generalise to phonotactic properties of words heard in the input to new words?

Hypothesis 10: There will be differences in participants' accuracy (only Dprime scores)[46] between *generalisable* and *non-generalisable distractors* in the word recognition task.

In particular, Dprime scores on generalisable distractors will be lower than that on non-generalisable distractors because participants will think that generalisable distractors are possible targets, and non-generalisable distractors are not.

Hypothesis 11: There will be no differences between *targets* and *generalisable distractors* in participants' performance in the forced-choice task.

In particular, performance will be similar between targets and generalisable distractors in the forced-choice task, given that both types of stimuli contain the same phonotactics. Participants will incorrectly think that generalisable distractors are targets.

(vi) Effect of linguistic training

11. (RQ) Is there a difference between linguistically sophisticated participants and linguistically-naïve participants in their ability to detect new words in Russian with respect to all hypothesis (1-12) formulated above.

[46] Dprime index of sensitivity is discussed in Section 6.1.

Hypothesis 12: There will be differences between *linguistically-naïve* and *linguistically sophisticated* participants in their accuracy/performance with respect to each hypothesis above (1-11).

In particular, linguistically-sophisticated participants are expected to have higher accuracy/performance than linguistically-naive participants with respect to each hypothesis (1-11).

5.5 Summary

Chapter 5 has provided a comprehensive review of the methodology of the present study. That is, participants, experimental stimuli, the procedure of the experiments and experimental tasks were described. These were followed by the formulation of research questions of the present study, and hypotheses bearing in mind conditions of the experimental design.

Chapter 6. Results

6.1 Introduction

In this chapter, the results of the word recognition and the forced-choice tasks are presented in Section 6.2 because the purpose of both tasks was to see if participants could detect words that they had heard in the input as opposed to words they did not hear. Section 6.2 starts with a comparison of what the word recognition and forced-choice tasks involved, and an explanation of why different statistical techniques were used for the analysis of performance in these tasks. This is followed by sections dedicated to the testing of specific hypotheses. These sections are titled to reflect the essence of what variables and interactions between variables are being investigated. Within each, the results from the linguistic group are discussed because the 12th research question concerned the performance of linguistically sophisticated as opposed to linguistically naïve participants concerning each hypothesis. A summary of the main results is given at the end of each section. The results for the cognate identification task are presented in Section 6.3, and a summary of the main results is presented in Section 6.4.

6.2 Results of the word recognition and forced-choice tasks

In the word recognition task, participants were asked to listen to 144 words, which were presented to them one by one, including 48 targets, 48 generalisable items (distractors), and 48 non-generalisable items (distractors). Participants were asked to decide whether or not they had heard these words previously in the input phase. The 48 targets were stimuli which had appeared in the input, and 48 generalisable items and 48 non-generalisable items were new, together equalling 96 new items (distractors), which did not appear in the input. Meanwhile, in the forced-choice task, participants were asked to listen to 48 pairs of experimental stimuli presented one by one, where one in the pair was a target – stimulus which had appeared in the input, and the other was a new generalisable item. The participants needed to decide which of the two items in each pair they had encountered in the input.

The word recognition task had an unbalanced design, with more new items (*n*=96) than old items (*n*=48), so that there were two equal groups of distractors (generalisable vs non-generalisable) in order to investigate the participants' generalisation of phonotactic abilities. However, this unbalanced design is problematic for a traditional statistical analysis with a binary dependent variable such as logistic regression because of the involvement of *response bias*. Response bias refers to a situation when participants show a tendency to give more "yes" responses (here where the target is present) than "no" responses (where no target is present) or *vice versa*. This situation can occur for a number of reasons, but one which is relevant to the present study is an unbalanced experimental design. More distractors than targets, as in the design of the word recognition task, can trigger *negative response bias*, which refers to a situation when some participants with a tendency to correctly reject words would achieve high accuracy because these participants would be correct for most of the distractors which constitute more items than targets. Consequently, for an unbalanced design, overall accuracy and response bias are confounded. Therefore, it was decided to use *Signal Detection Theory* (SDT) in the analysis for the word recognition task which overcomes this issue (Green & Sweets 1966). SDT is used in any discipline which involves a problem associated with decision making, as it specifically models response bias. To understand SDT, two fundamental concepts of a *signal* and *noise* must be understood. The signal is another name for a target, and noise is another name for a distractor. In a typical SDT task, two different types of stimuli, where one is the target and the other one is the distractor must be distinguished by a participant in order to measure the certainty of the ability to discriminate between these two types (Stanislaw & Todorov 1999). For example, in the word recognition task, participants needed to decide whether a word was present in the input phase or not; in other words, if this word was a target and present in the input, or a distractor and absent from the input. SDT presupposes four outcomes depending on a participant's responses, which for the analysis of the word recognition task were determined based on its design. It should be recalled that, in the word recognition task, participants needed to press the key “z” on the computer keyboard if they thought the word had been present, and to press “m” if it was absent. The four possible outcomes are:

(1) *Hit* = if the type of stimulus is “target”, and the response “z”;

(2) *Miss* = if the type of stimulus is “target”, and the response “m” or “none”;

(3) *False Alarm* = if the type of stimulus is “distractor”, and the response “z”;

(4) *Correct Rejection* = if the type of stimulus is “distractor”, and the response “m” or “none”.

Hits and correct rejections are indicative of *accurate responses*, whereas misses and false alarms are instances of *incorrect responses*. Based on these four outcomes, indices of sensitivity and discriminability, such as *d’, beta, A’, c*, were computed using *R* software (*R* Core Team 2013) using the *psycho* package (Makowski 2018). For the analysis of the word recognition task, the Dprime *(d’)* measure of sensitivity was used as it is a parametric measure of sensitivity which is probably the most commonly used among all such indices by researchers. What *d’* does is to measure the distance between the means of numbers of hits and false alarms in standard deviation units, and it is calculated using the following formula:

$$d' = z(\text{hit rate}) - z(\text{false alarm rate}),$$

where a hit is the presence of a target and the participant responds “yes”, and a false alarm is the presence of a distractor and the participant responds “no”, and z is the number of standard deviations from the mean (MacMillan & Creelman 2005).

According to Stanislaw and Todorov (1999), it can be difficult to interpret particular values of *d’* due to the use of standard deviations in its computation. However, the following interpretation of *d’* values is commonly used: a zero value of *d’* signals that participants can discriminate between a signal (target) and a noise (distractor) at a chance level (50%); and larger values signal good discriminability. For example, d'=4 signals excellent discriminability (at 100%); and negative values mean that participants performed below the level of chance. Figure 6-1 below was adapted from Azzopardi and Cowley (1998) to illustrate what *d'* prime values correspond to in terms of percentage accuracy on the y-axis, and number of hits on the x-axis[47]. The figure suggests that a value of *d'*=0.5 approximately corresponds to nearly 60% accuracy, *d'* values ranging from 0.5 to 1.5 correspond to accuracy rates from nearly to 60% to 75% respectively. It can be seen from the word recognition task analyses below that participants' responses varied from -0.10 to 1. Therefore, the limits on the y-axis for the word recognition task range from 0 to 1, or from -0.10 to 1 for those analyses where there are negative *d'* scores.

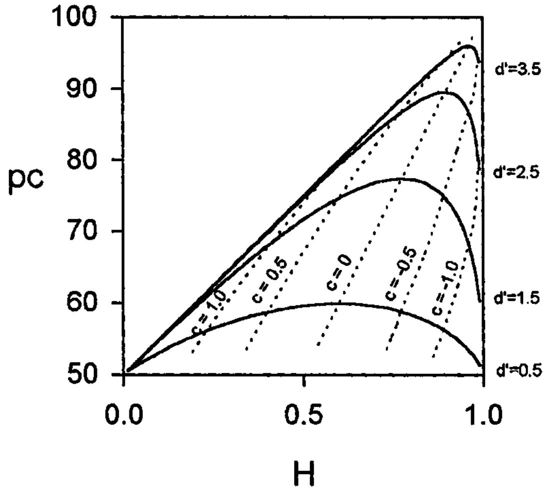

Figure 6-1. *d'* values corresponding to percentage correct and number of hits (pc=percent correct; H=hits) (adapted from Azzopardi and Cowley 1998: 295)

Unlike in the word recognition task, the forced-choice task utilised equal numbers of targets and distractors, and that is why a mixed-effect logistic regression model was used in the analysis of the results of the forced-choice task. A mixed-effect logistic regression model was chosen not only because of the balanced design of the forced-choice task, but also because it models the relationship between a response variable which is categorical. In this case, the response variable in the forced-choice task was binary (1 or 0, where "1" was correct, and "0" was incorrect) with one or more explanatory variables or predictors. Additionally, mixed-effects models are considered to be superior to traditional analyses based on quasi-F tests due to their ability to model variations according to random factors (that is by-subject and by-item), and random slopes (Baayen *et al.*

[47] Figure 6-1 shows the correspondence to percentage correct and number of hits in addition to *d'* value. The *c* value of correspondence may be of interest to the reader, but is not relevant for later discussion in this thesis.

2007). Finally, mixed-effect models are preferred as they can deal with the issue of missing values.

The structure of a mixed-effect logistic regression model closely resembles that of a linear regression model. However, unlike the latter, for which dependent variables should be on interval or ratio scales, a dependent variable in a mixed-effect logistic regression model must be binary. In this regression, a logit function is employed, and the outcome is expressed in log odds, the results of which are reported regarding *Odds Ratios* (OR). The OR is a measure which reflects the probability of one outcome (for example, the correct response "1") compared with the other outcome (for example, the incorrect response "0") for a specified predictor in a given model. An OR can be established for each predictor, and it expresses how the chances of a particular outcome change when the value of the predictor changes. After the mixed-effect logistic regression has been calculated, the exponential function of the regression coefficient is used to achieve the odds ratio associated with an increase in one unit of the explanatory variable. If OR=1, this means that the explanatory variable does not affect the outcome. If OR>1, it means that the chances of a correct outcome are greater than that of an incorrect outcome. And, if OR<1, it means that the chances of an incorrect response are greater than that of an incorrect outcome. The 95% *confidence interval* (CI) is used to estimate the precision of the ORs. A small CI range signals the higher precision of the OR, whereas a high CI range indicates lower precision. It is generally accepted if the accuracy of a model is statistically significant, the CI range should not include a zero. The measures, OR and CI, are used in reporting the results of the accuracy measure for the forced-choice task.

We now move to a description of how the analyses of the word recognition and forced-choice tasks were implemented. The analysis of accuracy for the word recognition task started by calculating the numbers of hits, misses, false alarms and correct rejections. Before each hypothesis can be tested, the data is aggregated and grouped anew according to those variables which are relevant to the specific hypothesis. This procedure produced new *d'* values for each hypothesis. One-way, two-way, and three-way repeated-measures *analyses of variance* (*ANOVA*s) were conducted. The *Dprime score* was always the dependent variable. Independent variables were the experimental conditions, such as: (a) *session* (session 1, session 2, session 3, vs session 4); (b) *length* (monosyllabic vs bisyllabic words); (c) *stress* (strong-weak vs weak-strong words); (d) *type of stimulus* (targets, generalisable distractors/items vs non-generalisable distractors); and (e) *phonotactic nativeness* (native vs non-native phonotactics). All the levels of data within these experimental conditions comprised of numbers of targets and distractors (see Section 5.2.1) which are necessary for the calculation of *d'* since the formula used to calculate it includes hits and false alarms. However, it should be recalled that per the design of the experiment, the *phonotactic nativeness* condition in addition to native phonotactics (which contained 24 targets and 48 distractors) and non-native phonotactics (which contained 24 targets and 24 distractors) included another level called "neither", which comprised of only 24 distractors without targets. The fact that the level "neither" did not contain corresponding targets was problematic for the calculation of *d'* because the formula for its calculation requires numbers of targets and distractors, and if one of these is missing, the calculation cannot proceed. Therefore, in order to determine the effect of the phonotactics variable with all three levels (native, non-native and neither), in addition to an ANOVA analysis with Dprime as a dependent variable and *phonotactic nativeness* (native vs non-native phonotactics) as an independent variable, a mixed-effect logistic regression model was fitted for accuracy (which was defined as number of hits and correct rejections) as a dependent variable, and phonotactic nativeness with all three levels as an independent variable. The mixed-effect logistic regression model for the phonotactic variable was not an ideal choice for the statistical analysis due to the unbalanced design of the phonotactics condition which entails that response bias will have a greater influence over the results than usual, for example someone with a negative response bias could achieve high accuracy because there are many distractors. However, as it was not possible to calculate *d'* for the level *neither* of the phonotactics variable, the mixed-effect logistic regression was the only statistical method available which could provide information about the effect of all levels of phonotactics although it is not designed to model response bias. Moreover, within the testing of each hypothesis testing, another analysis was conducted where in addition to dependent and within-subject independent variables (in ANOVA) and explanatory variables in the mixed-effect logistic regression analysis as discussed above, another between-subjects independent/explanatory variable was added to the analysis. This variable was (e) *linguistic group* (linguistically sophisticated participants vs linguistically naïve participants).

All statistical analyses in the present experiment were run in *R* (*R* Core Team 2013). All mixed-effect logistic regression models were conducted utilising the package *lme4* (Bates *et al*. 2013). All mixed-effect logistic regression models for the word recognition task, in addition to dependent and explanatory variables, included by-subject (*subject*) and by-item (*word*) random intercepts. For each hypothesis, a model with both random factors was compared with another model where the by-item (*word*) random factor was dropped. The comparison of models was accomplished using ANOVA. In all cases, the models with both random factors fitted were significantly better.

For the analysis of the word recognition task using ANOVA, two participants were excluded from the data analysis due to an Open Sesame technical fault logging their responses in session 3, as well as the inability of the package *afex* (Singmann *et al.* 2015) to deal with missing values for the whole analysis of the word recognition task. Last, but not least, the Dprime calculation does not model variation among items, and all results to be reported for the word recognition task in the ANOVA are by-subjects only.

As for the forced-choice task, the analysis of accuracy began by identifying whether participants performed correctly or incorrectly in each trial. Recall that, in the forced-choice task, participants needed to press the key “z” on the computer keyboard if they thought the first word had appeared in the input and to press “m” if they thought the second word had. To establish accuracy in this task, it was sufficient to know if the first word was a target or a foil, and how participants responded to it. A hit was counted when either of the following conditions was met: stimuli type "target" and response "correct", and stimuli type "distractor" and response "correct". A miss was counted when either of the following was met: stimulus type= "target" and response "incorrect" or "none", and stimuli type= "distractor" and response "incorrect" or "none". As a result of this procedure, the new variable of *accuracy* was coded as "1" if the response was correct and "0" if it was incorrect. Accuracy was a dependent variable used in the analysis of each hypothesis for the forced-choice task except for that concerning generalisation abilities (see Section 6.2.5). Explanatory variables were the same as independent variables in the word recognition task, i.e. *session*, *length*, *stress*, *type of stimuli*, *phonotactic nativeness*, and *linguistic group*. Due to the design of the forced-choice task[48], phonotactic nativeness variables consisted of only two levels (native vs non-native), and the type of stimulus variable included targets and generalisable distractors. For the forced-choice task, the effects on the detection of words were always assessed using the mixed-effect logistic regression models, also utilising the package *lme4* (Bates *et al.* 2013). For the forced-choice task, all models constructed for hypothesis testing, in addition to dependent and explanatory variables, included by-subject (*subject*) and by-item (*word1*) random intercepts. For each hypothesis, a model with both random factors was compared to another model where the by-item (word1) random factor was dropped. The comparison of the models was conducted in ANOVA. In all cases, models with both random factors fitted were significantly better, which indicates that retaining the random effects in the models is justified. Subsequent sections report the results of analyses by-participants and by-items. All the explanatory variables were factors which are the same as in the word recognition task.

[48] Design of the forced-choice task is explained in Section 5.2.

6.2.1 Detection over sessions

A one-way ANOVA was conducted with Dprime scores as a dependent variable and session number as an independent variable so as to investigate whether or not the detection of words from the input increases over sessions. Before the analysis, the data were aggregated according to Dprime score, subject and session variables.

The analysis of variance revealed significant differences for the main effect of session [$F(3, 75)=12.44$, $p< 0.001$, $\eta_p^2=0.33$][49]. A planned pairwise comparison showed a statistically significant increase in Dprime scores from sessions 1 to 2 [$F(3, 75)=12.44$, $p=0.05$], with the mean of *d'* on day 1 [$M=0.20$], and on day 2 [$M=0.36$]. There was no significant increase in scores from sessions 2 to 3 [$F(3, 75)=12.44$, $p=0.86$], with the mean of *d'* on day 3 [$M=0.40$]. Critically, there was a significant difference between Dprime scores on sessions 3 and 4 [$F(3, 75)=12.44$, $p=0.04$], with the mean of *d'* on session 4 [$M=0.57$] which was the highest score among all four sessions. Moreover, there was a statistically significant difference in Dprime scores between sessions 1 and 3, [$F(3, 75)= 12.44$, $p< 0.001$], and also the difference in Dprime scores was significant between sessions 2 and 4 [$F(3, 75)=12.44$, $p< 0.001$]. These results indicate significant improvements over time with only the session 2 versus session 3 comparison yielding a non-significant result. The value of Dprime=0.57 which participants scored on session 4 corresponds to a number slightly higher than 65% of accuracy but, importantly, what would be expected, the performance was above chance on all sessions, even on the first day. Figure 6-2 below represents the Dprime scores for each session for the word recognition task.

---

[49] Effect size is expressed by *partial eta-squared* (ηp2). This value expresses the amount of variance in the dependent variable as it was affected by an independent variable. According to Cohen's (1988) guidelines: (1) if the ηp2 values ≤0.01, the effect size is small; (2) if ≤0.059, the effect size is medium; and (3) if ≤0.138, the effect size is large. For example, if ηp2=0.33, this means that the effect size is very large, and that 33% of the change in the dependent variable, e.g. "Dprime score" as in the first hypothesis testing, can be explained by the independent variable "session". ηp2 is reported by default when using the *aov.car* function in the *affex* package in *R*.

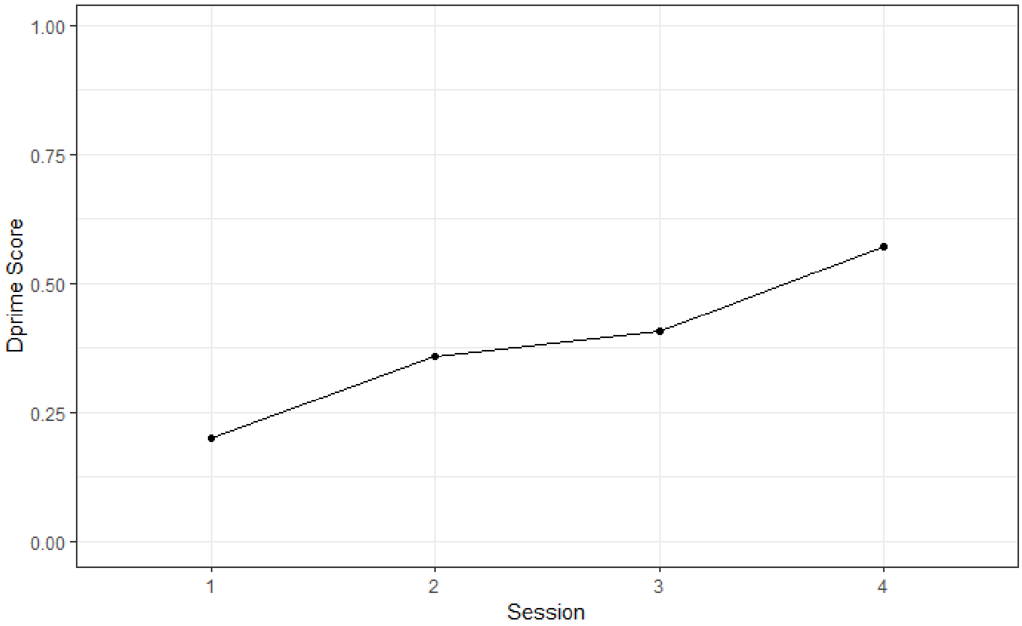

Figure 6-2. Word recognition task: Dprime scores for each session.

For the forced-choice task, a mixed-effect logistic regression model was constructed with accuracy as a dependent variable and session as an explanatory variable to investigate whether or not the detection of words from the input increased over sessions. The results for the model indicate a significant effect of session 4 [OR=1.4 (95% confidence interval [CI]: 1.19, 1.62), $p<0.01$], but the effect was not significant for session 2 [OR=1.10 (95% confidence interval [CI]: 0.94, 1.29), $p=0.21$], and was only marginally significant for session 3 [OR=1.14 (95% confidence interval [CI]: 0.97, 1.33), $p=0.09$]. These values indicate the changes in accuracy in comparison with the baseline level on the first day. These results show that ORs for all sessions are above one, which means that accuracy increased each day, and in particular, the mean percentage accuracy on session 4 was 56.7%, 52.1% on session 3, 51.3% on session 2, and 49% on session 1. This trend in increasing accuracy increase illustrated in Figure 6-3.

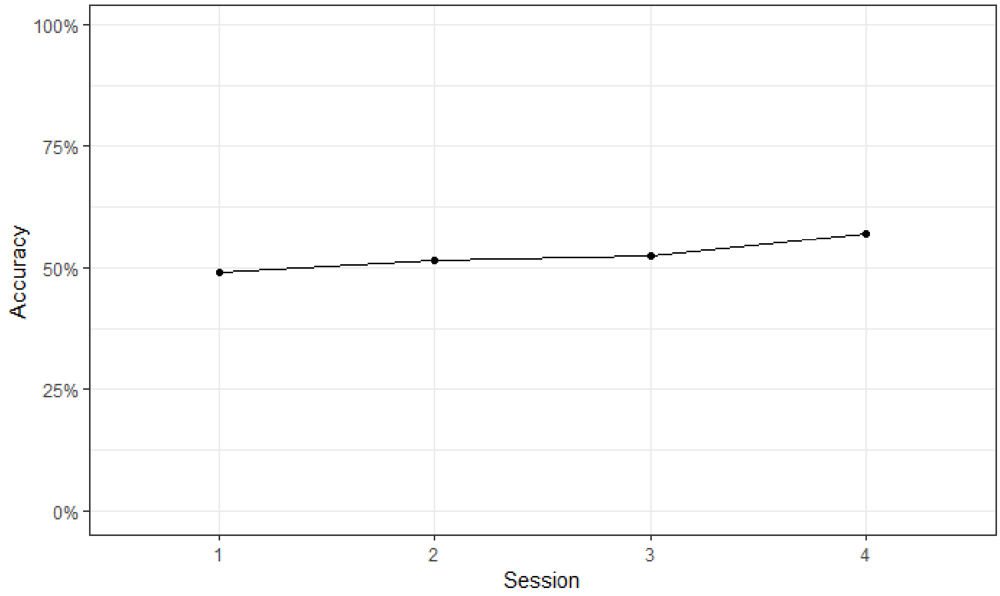

Figure 6-3. Forced-choice task: mean percentages of accuracy scores for each session.

6.2.1.1 Linguistic group and detection over sessions

For the word recognition task, a two-way mixed ANOVA with Dprime scores as a dependent variable, session as an independent within-subject variable, and linguistic group (sophisticated vs naïve) as a between-subject variable was conducted to investigate whether or not there are differences between linguistically sophisticated and linguistically naïve participants in their ability to detect words over four sessions. Before the analysis, the data were aggregated according to Dprime score, subject, session and group.

The analysis of variance revealed that the main effect of an interaction between sessions and linguistic group was not significant $F(3,75)=0.48$, $p=0.68$, $\eta_p^2=0.02$. However, despite this statistically insignificant main effect, it is evident from figure 6-4 below that the linguistically sophisticated group [M=0.40] performed slightly better than the linguistically naïve group in session 2 [M=0.32], just as the former group's performance was better on session 3, with means of Dprime scores [M=0.47] for the linguistically sophisticated group and [M=0.34] for the linguistically naïve participants. There was also a difference in session 4 in favour of the linguistically sophisticated group who received the higher mean of Dprime score [M=0.60], whereas the mean for the linguistically naïve group was [M=0.54].

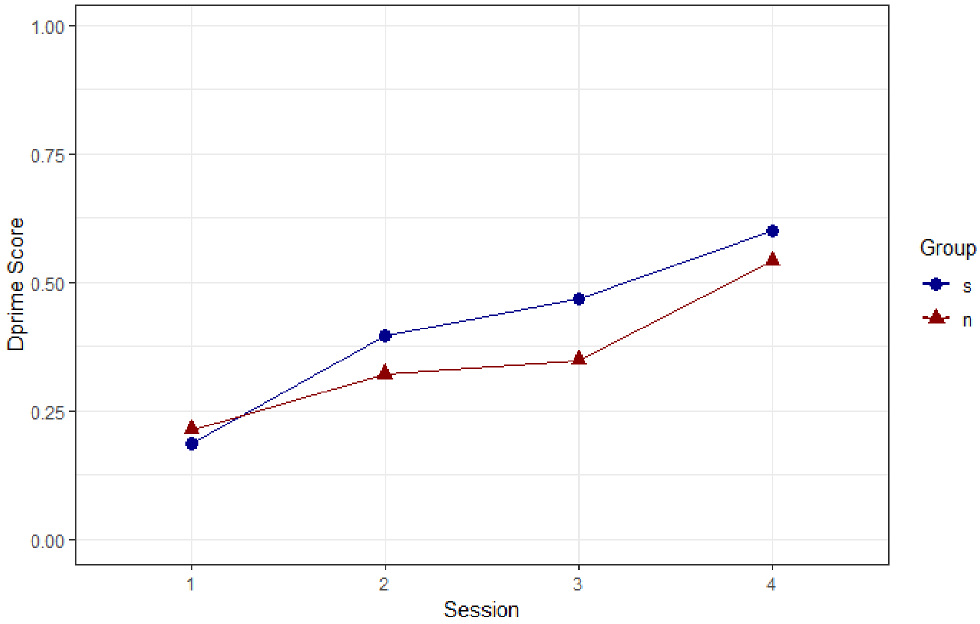

Figure 6-4. Word recognition task: Dprime scores in each session by linguistic group: s=sophisticated group; n=naïve group.

As for the forced-choice task, a mixed-effect logistic regression model was constructed with accuracy as a dependent variable and session and linguistic group as explanatory variables to investigate if there are differences between linguistically sophisticated and linguistically naïve participants in the ability to detect words over the four sessions. None of the comparisons within this model were significant; for the interaction between session 2 and naïve group [OR =1.02 (95% confidence interval [CI]: 0.75, 1.40), $p<0.88$], and an interaction between session 3 and naïve group [OR=0.97 (95% confidence interval [CI]: 0.71, 1.32), $p<0.84$], and the interaction between session 4 and naïve group [OR=0.19 (95% confidence interval [CI]: 0.59, 1.11), $p<0.19$]. An OR lower than 1 for the latter interaction between session 4 and naïve group indicates that the difference between sessions 4 and 1 was more pronounced for the linguistically sophisticated group. The values of mean percentage accuracy for the interaction of the session and group variables are presented in Table 6-1 below, and Figure 6-5 illustrates this interaction. It is evident from the figure and table that the responses improved on each day and the performance levels were similar between the groups except for the final session with sophisticated subjects having a higher mean accuracy [M=59.4%] than naïve participants [M=53.7%].

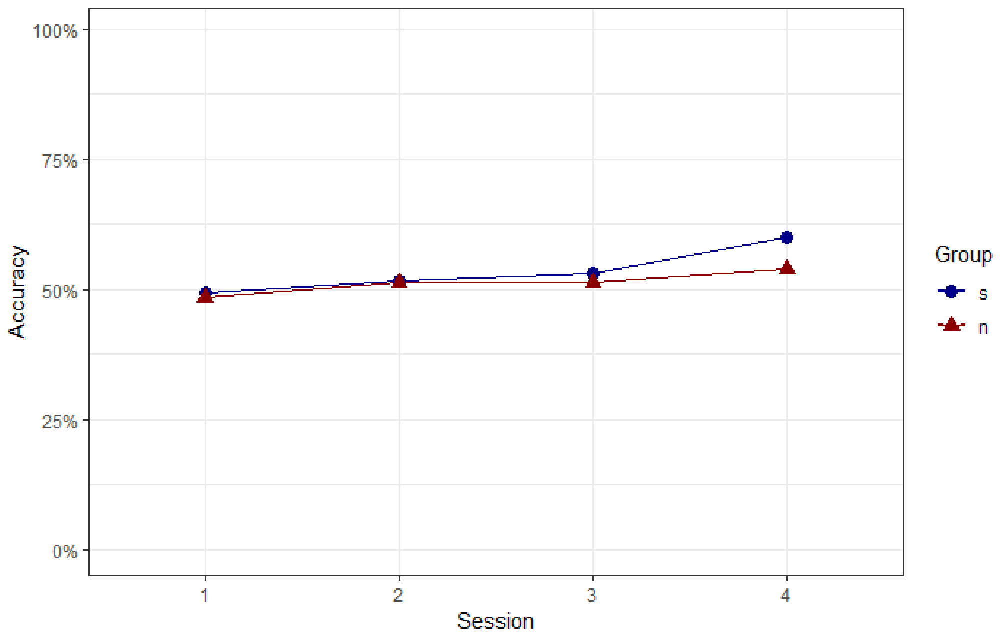

Figure 6-5. Forced-choice task: mean percentages of accuracy scores for each session by linguistic group: s=sophisticated group; n=naïve group.

Session	Linguistic group	
	Sophisticated	Naïve
1	49.4%	48.6%
2	51.5%	51.2%
3	52.9%	51.2%
4	59.4%	53.7%

Table 6-1. Forced-choice task: mean percentages of accuracy scores for each session by linguistic group.

6.2.1.1.1.1 Summary

To sum up, the results of the analysis of word recognition and forced-choice tasks found an effect of session number but to different extents. In particular, the results for the word recognition task showed that there was a significant difference in the ability to detect Russian words from the input in all sessions except between sessions 2 and 3. Meanwhile, the results from the forced-choice task showed only a marginally significant effect of session 3 and a significant effect of session 4. However, what was in common in the results for both tasks is that there was a clear trend of improvement throughout all sessions. These results confirm the *hypothesis 1* that an ability to detect words from the input would increase over sessions.

As for the linguistic groups, neither the results from the word recognition task nor those from the forced-choice task found any statistically significant differences between linguistically naïve and linguistically sophisticated participants. However, there were notable trends which showed that the linguistically sophisticated group was somewhat better at detecting words from the input over sessions. Furthermore, the results showed that, by the final session 4, the detection of target words was more pronounced for the linguistically sophisticated group. These results do not support *hypothesis 12,* which predicted that linguistically aware participants would perform better on word detection over sessions. The next section moves to the analysis of results, which demonstrates the cues which participants relied on for the detection of words.

6.2.2 Single cues

This section discusses the effects of each cue on the detection of words from the input. The section starts with phonotactics cues in Section 6.2.3.1, stress cues are considered in Section 6.2.2.2., and length cues in Section 6.2.2.3. Each of these sections is followed by the results for the effect of interaction between the linguistic group and each of the single cues.

6.2.2.1 Phonotactics

A one-way ANOVA was conducted with Dprime as a dependent variable and phonotactic nativeness (native vs non-native) as an independent variable in order to investigate whether or not the detection of words from the input depends on phonotactics. Before the analysis, the data were aggregated according to Dprime score, subject and phonotactic nativeness variables.

The analysis of variance revealed a significant main effect for the phonotactic nativeness condition [$F(1, 25)=14.64$, $p<0.001$, $\eta_p^2=0.37$]. This means that words with native phonotactics (i.e. those found in both English and Russian) were recognised significantly better [$M=0.35$] than words with non-native phonotactics (i.e. phonotactics found in only Russian) [$M=0.12$] which means that the effect is slightly higher than by chance. Figure 6-6 below illustrates the means of Dprime scores across native and non-native levels of the phonotactics condition.

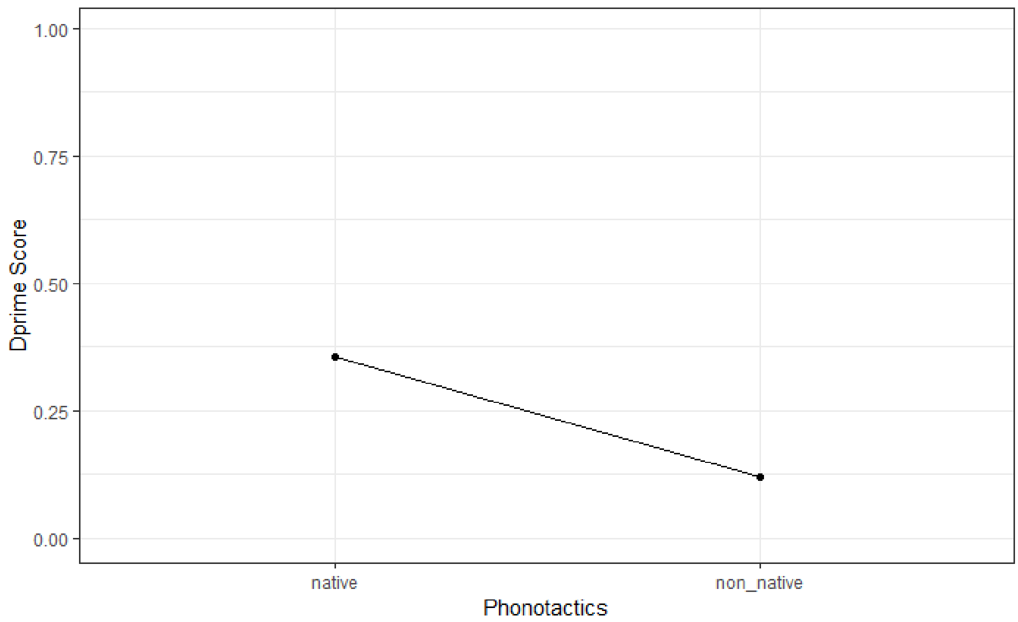

Figure 6-6. Word recognition task: Dprime scores for phonotactics condition.

However, it should be recalled from Section 6.1. that it was impossible to calculate Dprime scores for the phonotactics variable for the level *neither* because the experimental stimuli which represented this condition comprised only of distractors. To calculate Dprime scores, all levels of a variable or experimental condition should consist of some numbers of targets and distractors. As it was not possible to calculate this Dprime score, it was possible to run a mixed-effect logistic regression model on the phonotactics condition with all three levels (native vs non-native vs neither). However, due to the lack of modelling the response bias, the results from the mixed-effect logistic regression model should be interpreted with caution.

A mixed-effect logistic regression model was fitted where the dependent variable was accuracy which was binary (1="correct response", 0="incorrect response") so that a correct response was attributed to hits when participants identified targets correctly, and to correct rejections when participants knew that a distractor item was not a target. Hits indicate that participants could segment words while listening and match these representations to the target words at testing. Correct rejections indicate that participants could match the input against a set of stored target words from the input. The explanatory variable was the phonotactics condition with all three levels (native vs non-native vs neither). The results of the model indicated a significant effect of the non-native phonotactics condition [OR=0.77 (95% confidence interval [CI]: 0.64, 0.91), p<0.01]; and a significant effect of the neither phonotactic condition [OR=2.19 (95% confidence interval [CI]: 1.74, 2.76), p<0.01]. It is a rule of mixed-effect logistic regression that these values indicate a change in accuracy rates in comparison with the baseline level when phonotactics=native. This means that the OR with a value more than 1 for the phonotactics condition=neither, and a value less than 1 for the phonotactics condition=non-native, indicating that words which did not follow the phonotactics of either Russian or English were detected with the highest accuracy [*M*=75%] because all items in this experimental condition were distractors and participants were 75% correct to reject these words as not being targets. The accuracy for words with non-native phonotactics [*M*=58%] was significantly higher than that for words with non-native (i.e. Russian) phonotactics [*M*=52%], which was again just slightly above the chance level. The results of the mixed-effect logistic regression model for the word recognition task complement those from the ANOVA discussed at the beginning of this section. Figure 6-7 below illustrates the mean values of percentage accuracy in the phonotactics condition for the word recognition task.

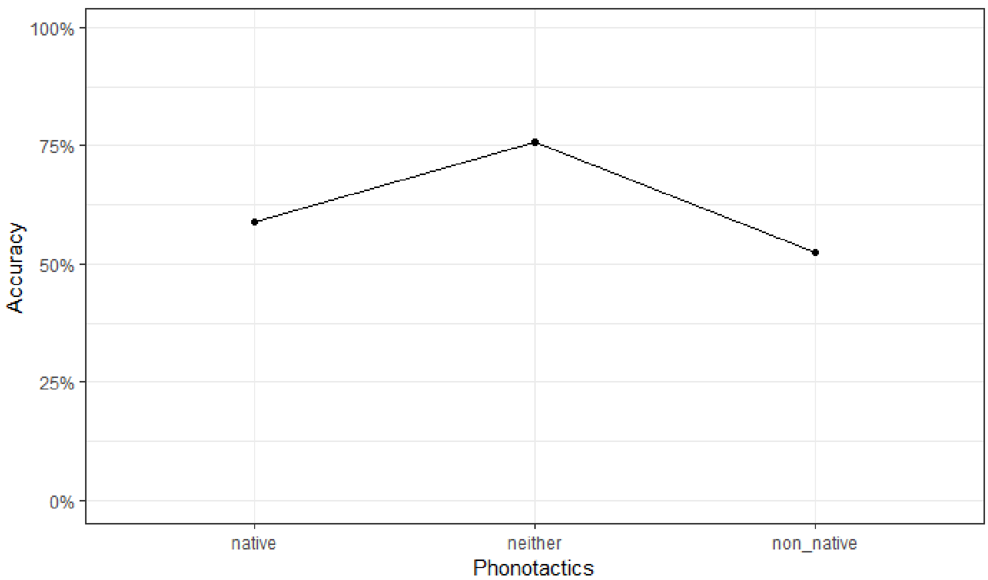

Figure 6-7. Word recognition task: mean percentages of accuracy scores for phonotactics condition.

As for the forced-choice task, a mixed-effect logistic regression model was constructed with accuracy as the dependent variable and phonotactic nativeness condition as an explanatory variable to investigate if there is an effect of native vs non-native phonotactics on the participants' ability to respond to words from the input. The results for the phonotactics condition variable for non-native phonotactics were not significant [OR=1.08 (95% confidence interval [CI]: 0.82, 1.45), p=0.57]. The slightly higher OR for the non-native phonotactics condition indicate that words with non-native phonotactics were detected slightly better [M=53.3%] over words with native phonotactics [M=51.2%]. Figure 6-8 below illustrates these results, which mean that type of phonotactics does not significantly affect the ability to respond to words from the input in the forced-choice task.

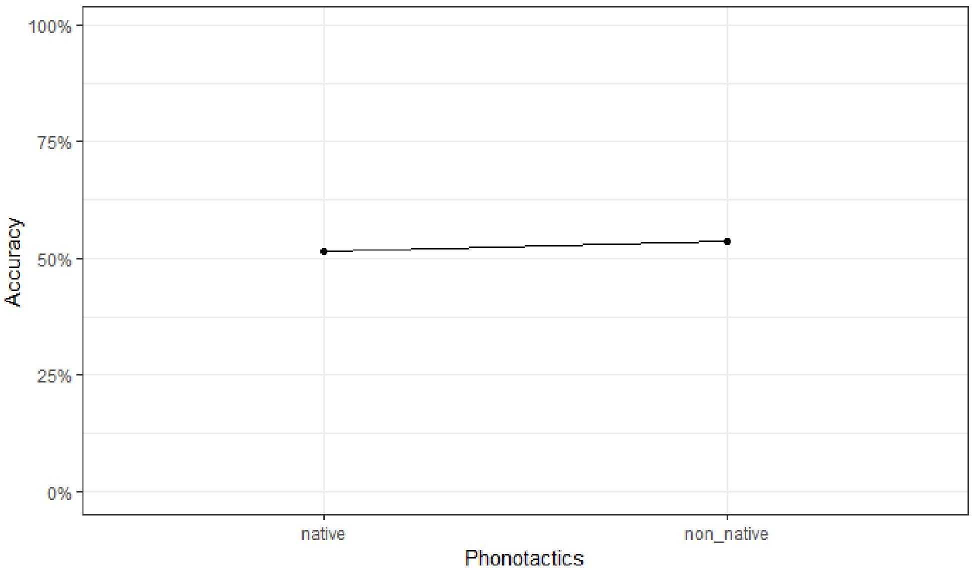

Figure 6-8. Forced-choice task: mean percentages of accuracy scores for phonotactics condition.

6.2.2.1.1 Linguistic group and phonotactics

A two-way ANOVA was conducted with Dprime as the dependent variable, phonotactic nativeness (native vs non-native) as a within-subject independent variable, and linguistic group (sophisticated vs naïve) as a between-subject independent variable to investigate whether or not the detection of words from input depends on interaction between phonotactics and linguistic group. Before the analysis, the data were aggregated according to Dprime score, subject, phonotactics, and linguistic group variables.

The analysis of variance revealed that this interaction was not significant [$F(1,24)=0.01$, $p=0.93$, $\eta_p^2=0.0004$]. Figure 6-9 below demonstrates that both linguistically sophisticated and naïve participants had higher accuracy in identifying words with the native phonotactic pattern ([$M=0.39$] for linguistically sophisticated, and [$M=0.32$] for linguistically naïve) than words with non-native phonotactics ([$M=0.16$] for linguistically sophisticated, and [$M=0.08$] for linguistically naïve). Additionally, a pairwise comparison of linguistic groups showed that linguistically sophisticated participants were better at identifying targets with native than non-native phonotactics [$p<0.05$]. The linguistically naïve group showed the same pattern [$p<0.05$]. However, once again, the main effect of the interaction of phonotactics and linguistic group was not statistically significant.

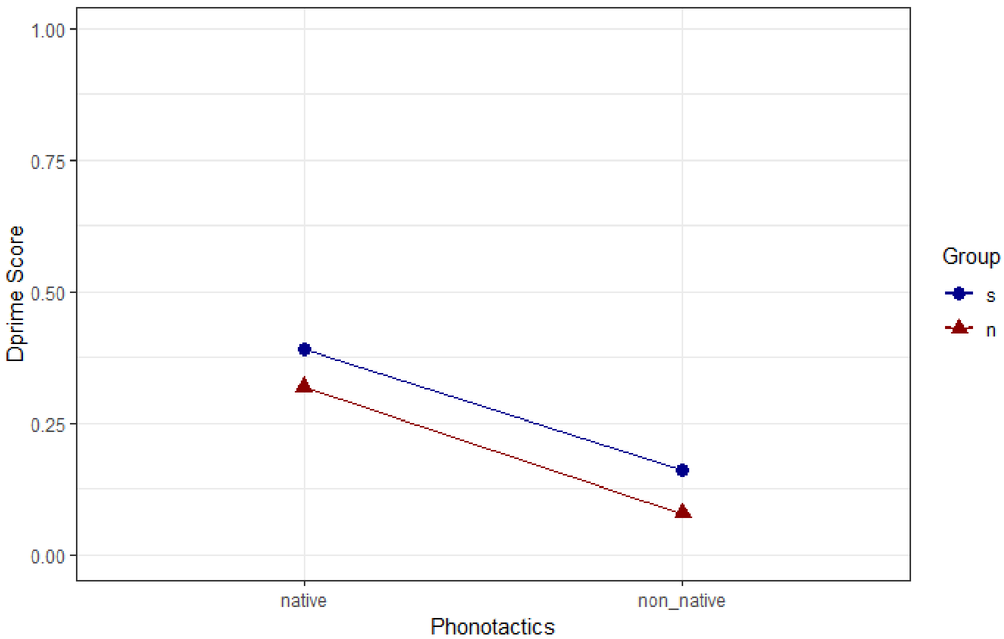

Figure 6-9. Word recognition task: Dprime scores for phonotactics condition by linguistic group: s=sophisticated group; n=naïve group.

Furthermore, to see the effect of all levels of the phonotactic condition, a mixed-effect logistic regression model was fitted where the dependent variable was accuracy rates for the phonotactics condition, and the explanatory variables were the phonotactics conditions with all three levels and linguistic group. The results of the model showed the following values for the interaction of phonotactics=neither and the sophisticated group [OR =1.07 (95% confidence interval [CI]: 0.87, 1.31), *p*=0.48] and the interaction of non-native phonotactics and sophisticated group [OR =0.88 (95% confidence interval [CI]: 0.76, 1.01), *p*<0.08]. Figure 6-10 illustrates this interaction. These results suggest that sophisticated participants were more accurate than linguistically naïve participants on all phonotactics conditions, but the general pattern was the same, so that words with phonotactics=neither received the highest accuracy, followed by words with phonotactics=native, and then performance on words with phonotactics=non-native for which accuracy was the lowest. The mean values of percentage accuracy for the interaction between phonotactics condition and linguistic group are presented in Table 6-2 below.

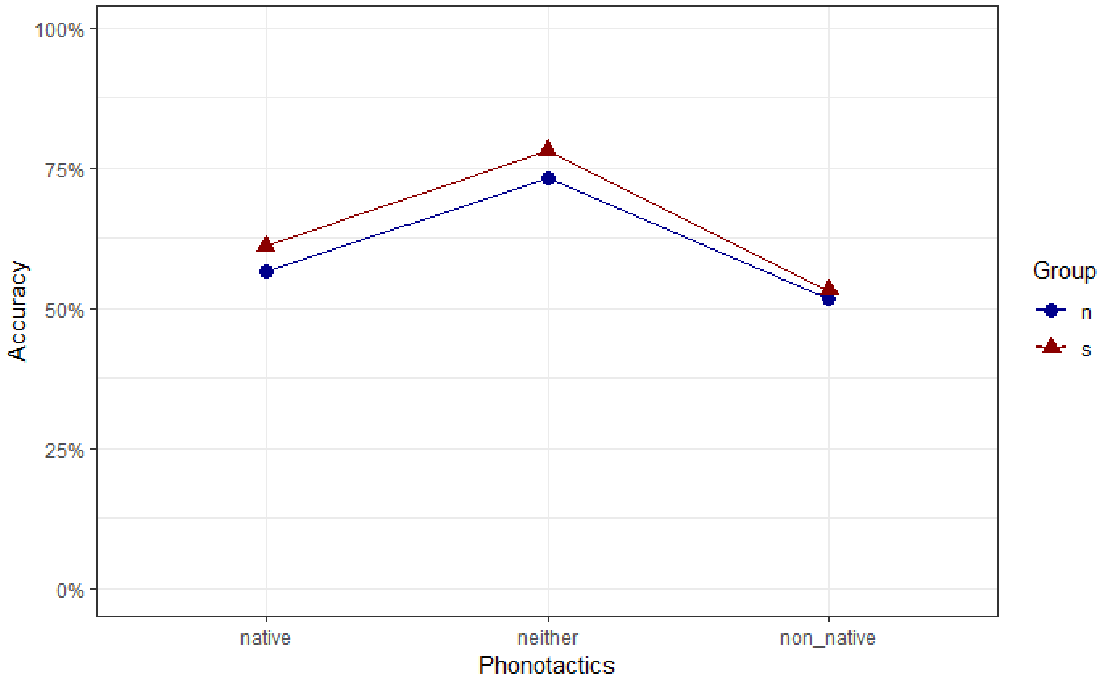

Figure 6-10. Word recognition task: mean percentages of accuracy scores for phonotactics condition by linguistic group: s=sophisticated group; n=naïve group.

Phonotactics	Linguistic group	
	Sophisticated	Naïve
Native	60.3%	56%
Non-native	52.9%	51.5%
Neither	77.1%	72.4%

Table 6-2. Word recognition task: mean percentages of accuracy scores on phonotactics condition by linguistic group.

For the forced-choice task results, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and phonotactics (native vs non-native) and linguistic group (linguistically sophisticated vs linguistically naïve) as explanatory variables in order to investigate whether or not the ability to respond to novel words is affected by the interaction between phonotactics and group. The results of the model show the following values for the interaction of non-native phonotactics and naïve group [*OR* =1.25 (95% confidence interval [*CI*]: 1.01, 1.57), $p<0.05$]. This suggests that there were statistically significant differences between non-native and native phonotactics for linguistically naïve participants, but these differences were nearly non-existent in linguistically sophisticated participants. Figure 6-11 below illustrates this interaction. The figure shows that linguistically naïve participants were more accurate in recognising words with non-native phonotactics [*M*=54%] than native phonotactics [*M*=47%]. Meanwhile, linguistically sophisticated participants recognised words with native phonotactics [*M*=53.5%] with nearly the same accuracy as words with non-native phonotactics [*M*=53.1%].

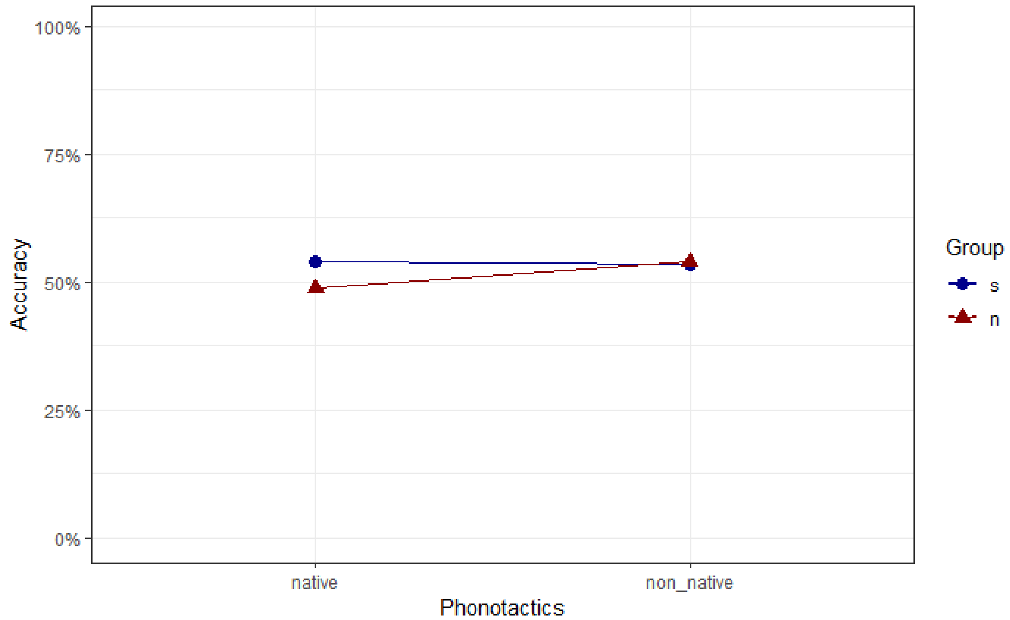

Figure 6-11. Forced-choice task: mean percentages of accuracy scores for phonotactics condition by linguistic group: s=sophisticated group; n=naïve group.

6.2.2.1.1.1 Summary

To sum up, the results for the word recognition and the forced-choice tasks showed different trends concerning the effects of phonotactics condition. In particular, the results for the word recognition task from ANOVA and mixed-effect logistic regression analysis showed that the words with native phonotactics were recognised better than words with non-native phonotactics. This supports *hypothesis 2,* which predicted that accuracy for words with native phonotactics would be higher than with non-native phonotactics. The results for the forced-choice task were statistically non-significant, with a minor trend indicating that words with non-native phonotactics were slightly preferred to words with native phonotactics. This means that the results from the forced-choice task do not support hypothesis 2.

As for linguistic group, there was no statistically significant difference between the linguistically sophisticated and naïve groups in the word recognition task. However, the results of the forced-choice task showed that linguistically sophisticated participants achieved statistically significantly higher accuracy (slightly above chance levels on words with native phonotactics) whereas the naïve group was less accurate on the native phonotactics condition (which performing below chance levels). Therefore, the results from the forced-choice task, but not from the word recognition task, support *hypothesis 12,* which predicted that there would be differences between these groups.

6.2.2.2 Stress

A one-way ANOVA was conducted with Dprime as the dependent variable and stress (strong-weak, and weak-strong) as the independent variable to investigate whether or not the detection of novel words depends on the stress patterns. Before the analysis, the data were aggregated according to Dprime score, subject and stress variables.

The effect of stress was not significant [$F(1, 25)=0.44$, $p=.51$, $\eta_p^2=0.02$], with the mean of Dprime scores for words with the weak-strong pattern [$M=0.30$], and [$M=0.25$] for words with the strong-weak pattern. s mean percentages of accuracy scores suggests that although the results were not significant, words with the weak-strong stress pattern were slightly preferred over words with the strong-weak pattern. Figure 6-12 below illustrates the means of Dprime scores across stress patterns.

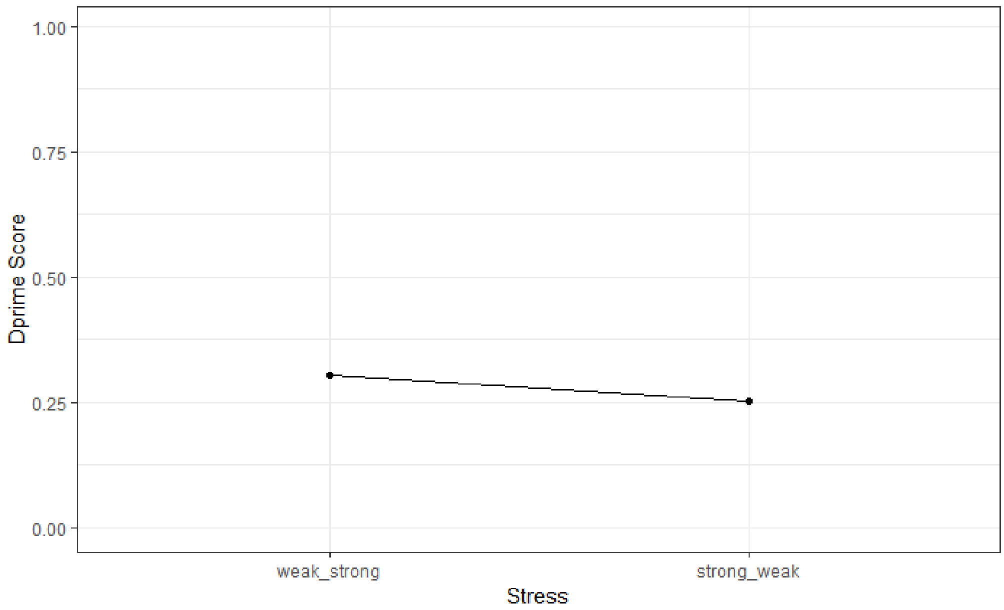

Figure 6-12. Word recognition task: Dprime scores for stress condition.

For the forced-choice task results, a mixed-effect logistic regression model was conducted with accuracy as the dependent variable and stress (strong-weak and weak-strong) as the explanatory variable to investigate if there is any effect of stress on the participants' ability to detect new words. The results showed a marginally significant effect of the strong-weak stress pattern [OR=0.70 (95% confidence interval [CI]: 0.48, 1.01), p=0.06]. The ORs which values less than 1 indicate that accuracy on words with strong-weak pattern was at 48.6%, which is lower than 57.2% the value of accuracy for words with the weak-strong pattern. Figure 6-13 below illustrates this pattern.

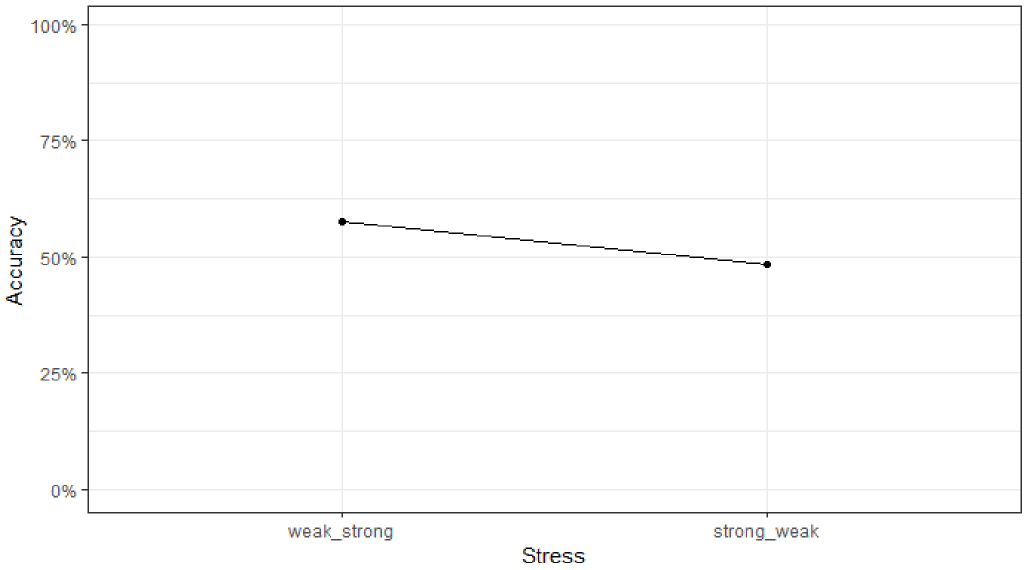

Figure 6-13. Forced-choice task: mean percentages of accuracy scores for stress condition.

6.2.2.2.1 Linguistic group and stress

A two-way ANOVA was conducted with Dprime as the dependent variable, stress (strong-weak vs weak-strong words) as a within-subject independent variable, and linguistic group (sophisticated vs naïve) as a between-subjects independent variable to investigate whether or not there are differences between linguistic groups in their detection of words from the input when relying on the stress patterns of words. Before the analysis, the data were aggregated according to Dprime score, subject, syllable stress, and linguistic group variables. The results of this interaction were not significant $F(1,24)=0.38$, $p=0.54$, $\eta_p^2=0.02$. Figure 6-14 below indicate that the linguistically sophisticated group performed almost the same on weak-strong words [$M=0.33$] and strong-weak words [$M=0.33$], whereas the naïve participants responded better to words with the weak-strong pattern [$M=0.26$] than words with the strong-weak pattern [$M=0.17$]. The comparisons were not significant at [$p=0.98$] for the sophisticated group, and [$p=0.37$] for naïve participants.

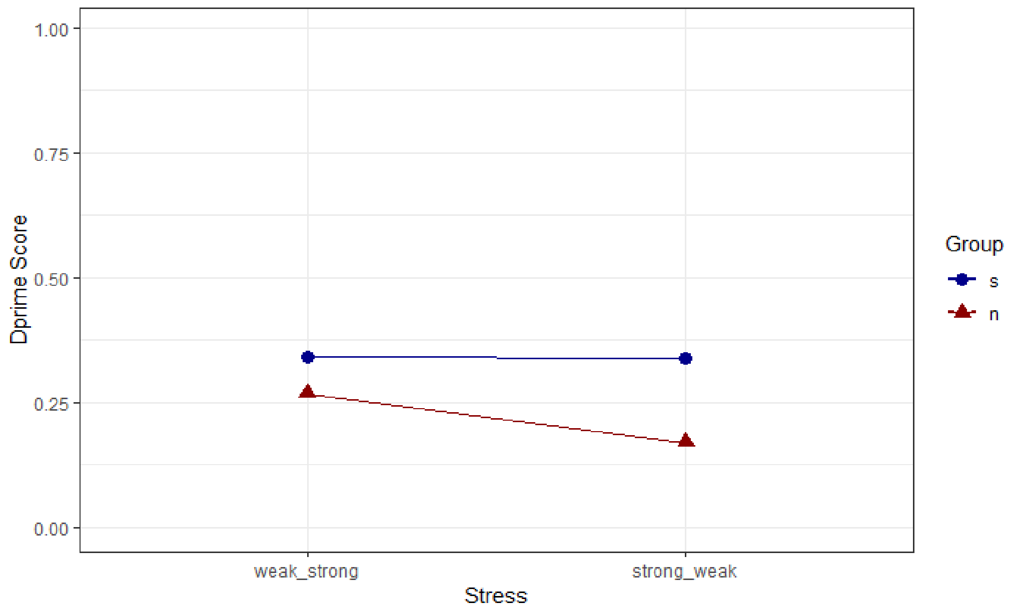

Figure 6-14. Word recognition task: Dprime scores for stress condition by linguistic group: s=sophisticated group; n=naïve group.

For the forced-choice task, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and stress (strong-weak vs weak-strong words) and linguistic group (linguistically sophisticated vs linguistically naïve) as explanatory variables to investigate whether there are differences between linguistic groups in their detection of words when relying on syllable stress. The results from the model showed that the interaction between the naïve group and words with the strong-weak pattern was not significant [OR =1.14 (95% confidence interval [CI]: 0.87, 1.50), p<0.33]. Figure 6-15 below illustrates this interaction. It shows that, for the naïve group, words with a weak-strong [M=55%] stress pattern were recognised slightly better than words with strong-weak stress [M=48%]. Similarly, the linguistically sophisticated group's accuracy was better on words with the weak-strong stress pattern [M=0.59] than words with the strong-weak pattern [M=49%].

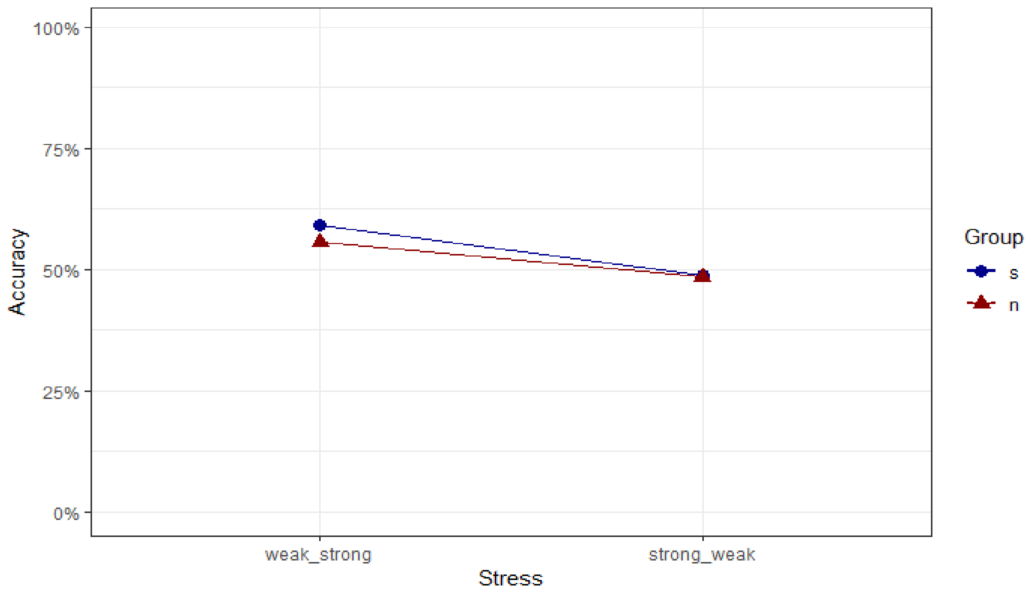

Figure 6-15. Forced-choice task: mean percentages of accuracy scores for stress condition by linguistic group: s=sophisticated group; n=naïve group.

Syllable Structure	Linguistic Group	
	Sophisticated	Naïve
Weak-strong	59%	55%
Strong-weak	49%	48%

Table 6-3. Forced-choice task: mean percentages of accuracy scores on stress condition by linguistic group.

6.2.2.2.1.1 Summary

The results from the forced-choice task exhibited marginal statistically significant difference in the sense that words stressed on the final syllable were detected better than words stressed on the first syllable, which were recognised nearly at chance levels. Also, the same trend was found in the word recognition task, although the difference was not statistically significant. This means that *hypothesis 3* should be rejected, since it predicted that words stressed on the first syllable would be detected better than words with final syllable stress.

There was no statistically significant difference between lingustic groups in their word detection abilities when relying on stress cues in both tasks. This once again does not support *hypothesis 12,* which predicted that linguistics students should perform better. However, the detection abilities of the linguistically sophisticated group were slightly higher than those of the naïve group.

6.2.2.3 Length

For the word recognition task, a one-way ANOVA was conducted to investigate the effect of the length of syllables within a word, where Dprime was the dependent variable, and length (monosyllables vs bisyllables) was an explanatory variable. The data were aggregated according to Dprime score, subject and length. The results showed no significant effect of syllable length [$F(1, 25)= 0.28$, $p=0.60$, $\eta_p^2=0.01$], with the Dprime score means of [$M=0.25$] for monosyllabic words and [$M=0.28$] for bisyllabic words. The results suggest that, despite the lack of a significant difference between the levels of the syllable length variable, bisyllabic words were slightly preferred over monosyllabic words. Figure 6-16 below illustrates the means of the Dprime scores across the syllable length condition.

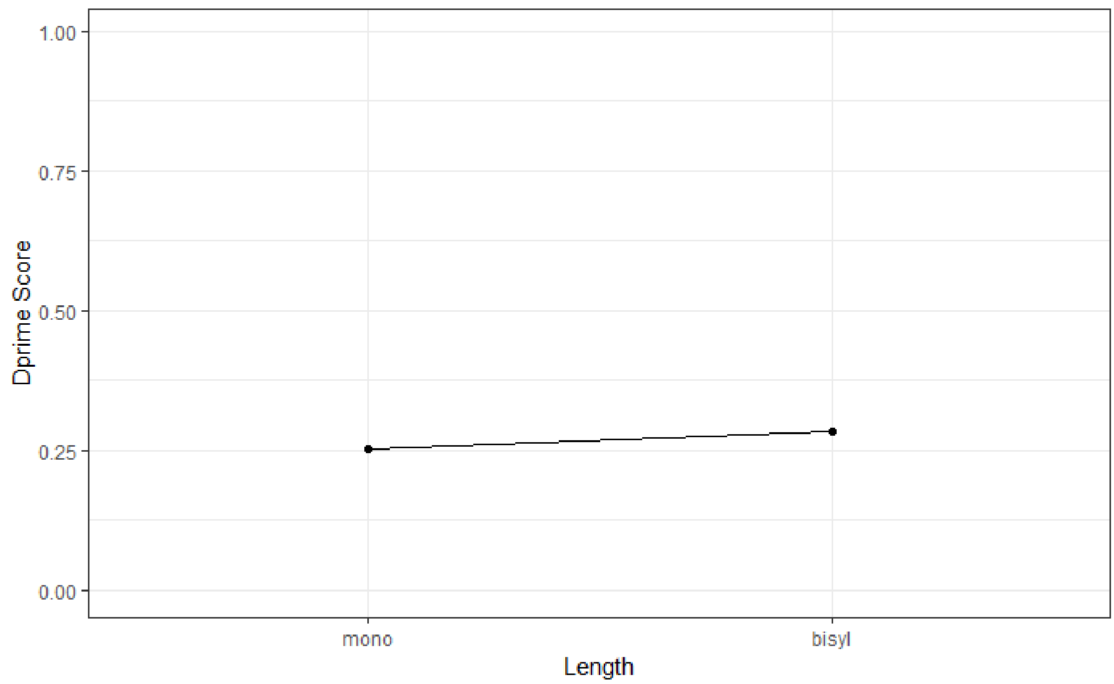

Figure 6-16. Word recognition task: Dprime scores for word length condition: mono=monosyllabic words; bisyl=bisyllabic words.

For the forced-choice task, a mixed-effect logistic regression model was fitted which contained accuracy as the dependent variable and length (monosyllabic vs bisyllabic words) as an explanatory variable. The results showed that there was no statistically significant effect of word length, with the following results [OR=1.07, (95% confidence interval [CI]: 0.79, 1.46), p=0.64], The mean accuracy was 51.2% for monosyllables and 52.8% for bisyllables. Figure 6-17 below illustrates these results.

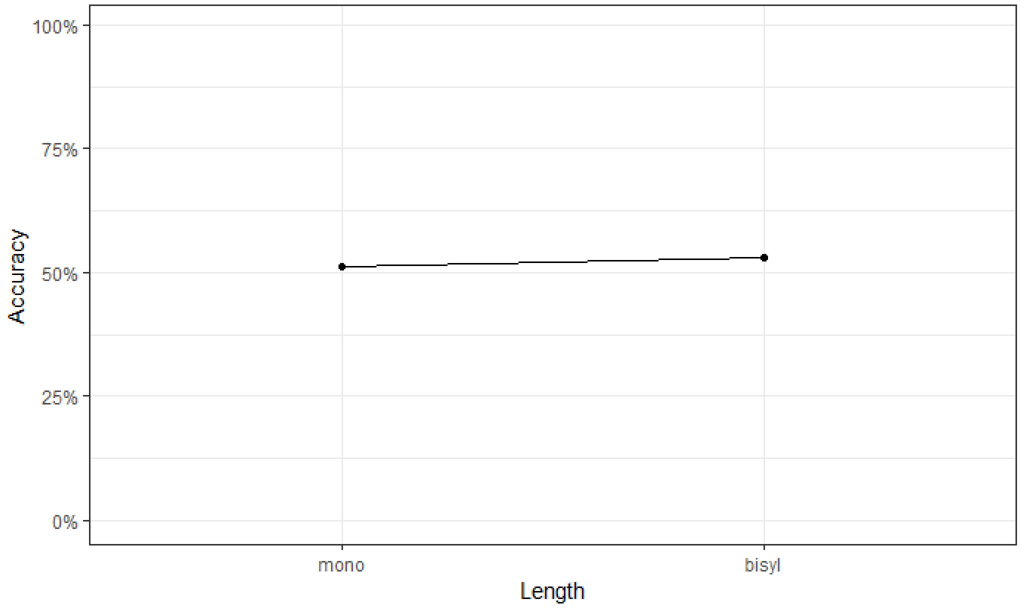

Figure 6-17. Forced-choice task: mean percentages of accuracy scores for word length condition: mono=monosyllabic words; bisyl=bisyllabic words.

6.2.2.3.1 Linguistic group and syllable length

For the word recognition task, a two-way ANOVA was conducted with Dprime as the dependent variable, word length (monosyllabic words vs bisyllabic) as a within-subject independent variable, and linguistic group (sophisticated vs naïve) as a between-subject independent variable in order to investigate whether or not there are differences between linguistic groups in their detection of words when relying on word length. Before the analysis, the data were aggregated according to Dprime score, subject, word length, and linguistic group variables. The results showed that the main effect of this interaction was not significant [$F(1,24)=1.23$, $p=0.28$, $\eta_p^2=0.05$].

Figure 6-18 below illustrates that, for the linguistically sophisticated group, bisyllabic words [M=0.34] were recognised better than monosyllabic words [M=0.25], whereas the means of Dprime scores for linguistically naïve subjects were slightly higher for monosyllabic words

[M=0.25] than bisyllabic ones [M=0.22]. Posthoc analysis showed that none of these comparisons were statistically significant [p=0.25], and [p=0.68] respectively.

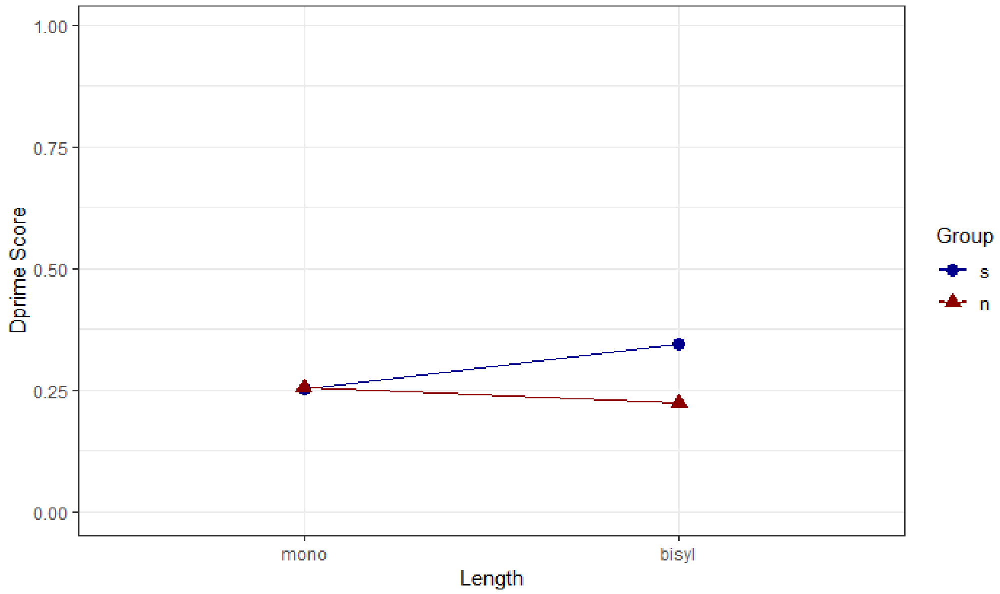

Figure 6-18. Word recognition task: Dprime scores on word length condition by linguistic group: s=sophisticated group; n=naïve group; mono=monosyllabic words; bisyl=bisyllabic words.

For the forced-choice task, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and word length (monosyllabic vs bisyllabic words) and linguistic group (linguistically sophisticated vs linguistically naïve) as explanatory variables to investigate if there are differences between the linguistic groups in their detection of words when relying on word length. The model's results showed that the interaction between naïve group and bisyllabic words was not significant [OR =1.03 (95% confidence interval [CI]: 0.82, 1.30), p<0.79]. This interaction is illustrated in figure 6-19 below, suggesting that both groups were slightly more accurate with bisyllabic words ([M=54%] for the sophisticated group, and [M=52%] for the linguistically naïve group) than monosyllabic words ([M=52%] for the sophisticated group, and [M=50%] for the linguistically naïve group), although the overall performance of the linguistically sophisticated group was slightly better.

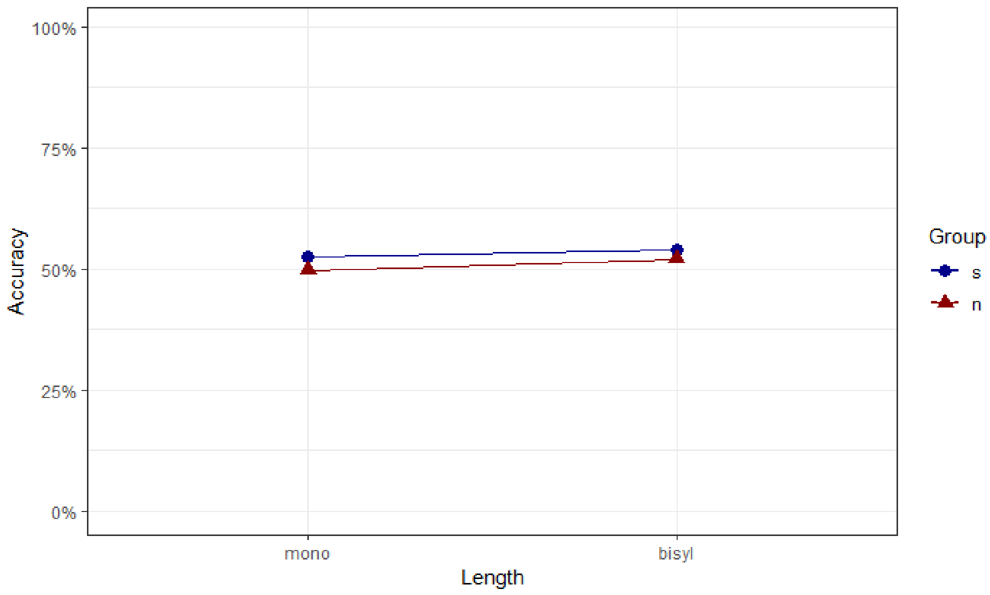

Figure 6-19. Forced-choice task: mean percentages of accuracy scores for word length condition by linguistic group: mono=monosyllabic words; bisyl=bisyllabic words; s=sophisticated group; n=naïve group.

6.2.2.3.1.1 Summary

The results from both tasks concerning the effect of word length on participants' ability to detect words from the input were not significant. These results do not support *hypothesis 4,* which predicted that bisyllabic words would be recognised more accurately than monosyllabic words. However, both tasks showed a trend for bisyllabic words to be recognised slightly better than monosyllabic words.

As far as the performance of the linguistic groups is concerned, there was no statistically significant difference in their word detection abilities when relying on word length cues in both tasks. This once again does not support *hypothesis 12,* which predicted that linguistic students should perform better. However, notably, the detection abilities of the linguistically sophisticated group were slightly better than those of the naïve group. The next section describes the results concerning interactions between phonotactics and stress pattern, and phonotactics and word length.

6.2.3 Interaction of cues

This section discusses the effects of the interaction between cues on word detection. The interaction between phonotactics and stress is first considered in Section 6.2.3.1, followed by that between phonotactics and word length in Section 6.2.3.2. Each of these sections includes a discussion of any combined effect of linguistic group and each of these interactions. The potential interaction between word length and stress is not investigated because such an interaction is not in the present study logically since the stress variable has only two levels (strong-weak and weak-strong) and operates only on bisyllabic words.

6.2.3.1 Interaction of phonotactics with stress

For the word recognition task, a two-way ANOVA was conducted with Dprime as the dependent variable, and phonotactics (native vs non-native) and stress (strong-weak vs weak-strong) as independent variables so as to investigate if the ability to respond to novel words depends on an interaction of phonotactics and stress cues. Before the analysis, the data were aggregated according to Dprime score, subject, phonotactics and stress. The results showed that the interaction between stress and phonotactics had a statistically significant effect [$F(1, 25)=16.47$, $p<0.001$, $\eta_p^2=0.40$]. Posthoc analysis showed that weak-strong words with the native phonotactic pattern were recognised significantly better [M=0.53] than weak-strong words with non-native phonotactics [M= -0.04], $p< .05$; whereas the difference in performance for strong-weak words with native phonotactic pattern and non-native phonotactic pattern was not statistically significant [p=1.0]. This is because the mean Dprime values for strong-weak words and for native phonotactics [M=0.26], and non-native phonotactics [M=0.26] were identical. Additionally, when the phonotactics were non-native, words with the strong-weak stress pattern [M=0.26] were recognised significantly better [$p<0.05$] than words with the weak-strong stress pattern [M= -0.04]. In contrast, when the phonotactics were native, words with the strong-weak stress pattern [M=0.26] were recognised significantly les soften [$p<0.05$] than words with the weak-strong stress pattern [M=0.53]. Table 6-4 below demonstrates the means of Dprime scores for the interaction between phonotactics and stress, and Figure 6-20 illustrates this interaction. The results suggest that weak-strong words with native phonotactics were detected with the highest accuracy, and weak-strong words with non-native phonotactics the lowest accuracy, whereas there was no difference in the recognition of strong-weak words across the phonotactics condition.

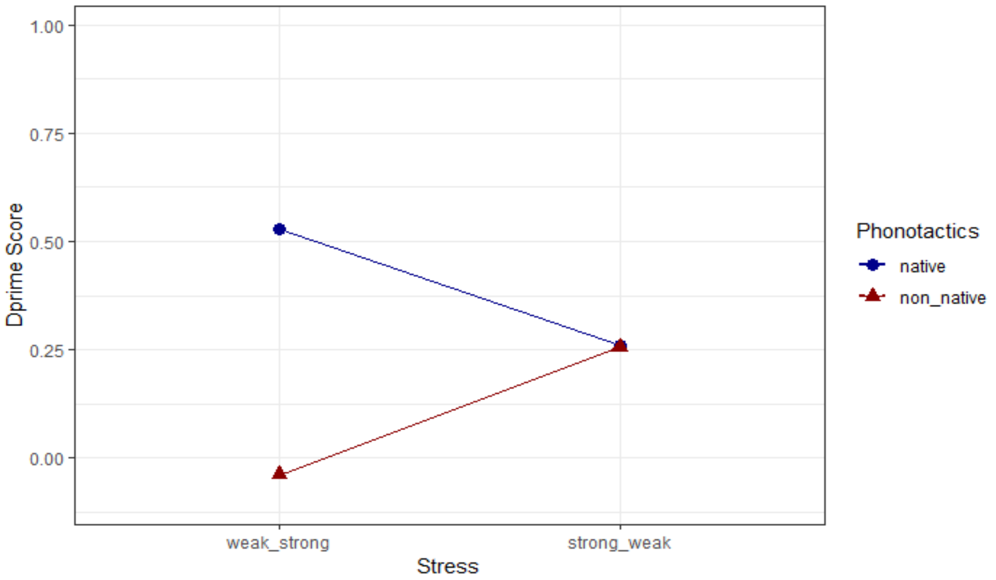

Figure 6-20. Word recognition task: Dprime scores for interaction of phonotactics and stress conditions.

Phonotactics	Stress	
	Weak-strong	Strong-weak
Native	0.53	0.26
Non-native	-0.04	0.26

Table 6-4. Word recognition task: Dprime scores for interaction of phonotactics and stress conditions.

Furthermore, to incorporate all levels of the phonotactics variable, a mixed-effect logistic regression model was conducted where the dependent variable was accuracy for a phonotactic condition, and the explanatory variables were all three levels of phonotactics and stress (strong-weak vs weak-strong). In the resulting model of the model the effect of the interaction between words with non-native phonotactics and strong-weak stress was marginally significant [OR=1.48 (95% confidence interval [CI]: 0.97, 2.27), p=0.07], and not significant between words with with the phonotactics of neither language and strong-weak words [OR=1.11 (95% confidence interval

[*CI*]: 0.64, 1.92), *p*=0.71]. What these results suggest, despite, strictly speaking, statistically insignificant results, is that words with phonotactics of neither language regardless of the stress pattern, at [*M*=76.8%] for weak-strong, and [*M*=75.3%] for strong-weak, were recognised at the highest accuracy. The next strongest effect was for words with native phonotactics, where words with weak-strong stress [*M*=60.3%] scored higher accuracy than words with weak-strong stress [*M*=55.9%]. However, words with non-native phonotactics were associated with the highest accuracy when the stress pattern was strong-weak [*M*=54.6%], while accuracy for weak-strong words and non-native phonotactics was the same as by chance [*M*=49.5%]. These results, despite statistical insignificance, complement those from the ANOVA discussed above, showing that words with non-native phonotactics and weak-strong stress pattern were most accurately detected. The mean percentage accuracy scores for the interaction between phonotactics and stress conditions are illustrated in Figure 6-21, while Table 6-5 summarises the values of mean accuracy.

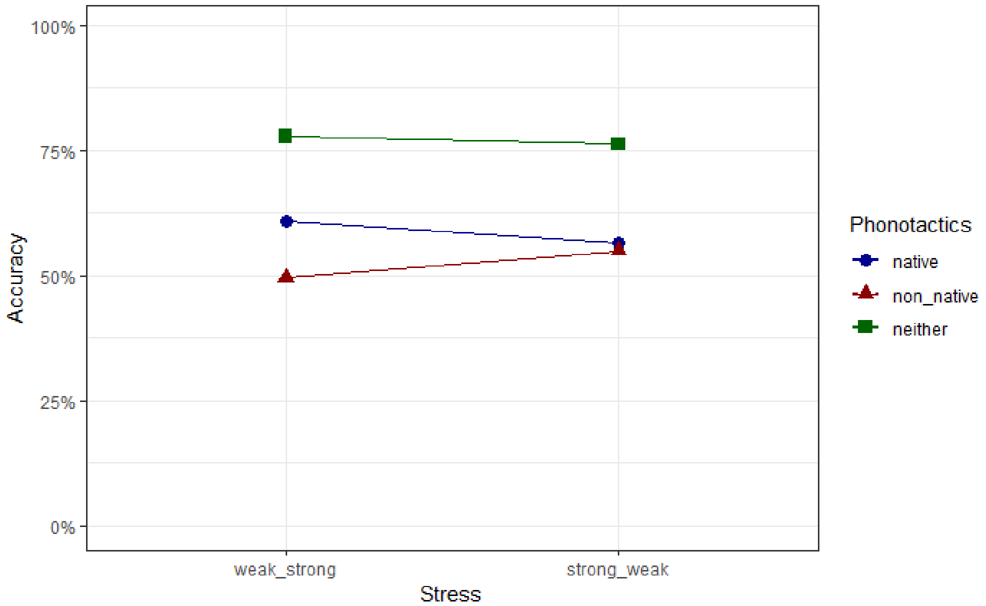

Figure 6-21. Word recognition task: mean percentages of accuracy scores for interaction between phonotactics and stress conditions.

Phonotactics	Stress	
	Weak-strong	Strong-weak
Native	60.4%	55.9%
Non-native	49.5%	54.6%
Neither	76.8%	75.2%

Table 6-5. Word recognition task: mean percentages of accuracy scores on interaction between phonotactics and stress conditions.

For the forced-choice task, a mixed-effect logistic regression model was fitted where accuracy was the dependent variable, and phonotactics and stress were explanatory variables in order to investigate if there was any effect of interaction between phonotactics and stress on participants' ability to respond to words from the input. The results of this model indicate a significant interaction between non-native phonotactics and words with strong-weak stress patterns, [OR=2.49 (95% confidence interval [CI]: 1.25, 4.95), $p<0.01$]. That is, values of mean percentage accuracy across the phonotactics condition, the mean percentages of accuracy were highest for words with the weak-strong words and native phonotactics [M=62.6%]. Whereas, words with the strong-weak stress pattern and native phonotactics received the lowest values of mean percentage accuracy. However, there was little difference between weak-strong words and strong-weak words when these words comprised of non-native phonotactics, with values of mean percentage accuracy of 51.9% for weak-strong, and 54% for strong-weak words. Table 6-6 below summarises the mean percentages of accuracy for the interaction of stress and phonotactics condition, and Figure 6-22 illustrates this interaction.

Phonotactics	Stress	
	Weak-strong	Strong-weak
Native	62.6%	43%
Non-native	51.9%	54%

Table 6-6. Forced-choice task: mean percentages of accuracy scores for interaction of phonotactics and stress condition.

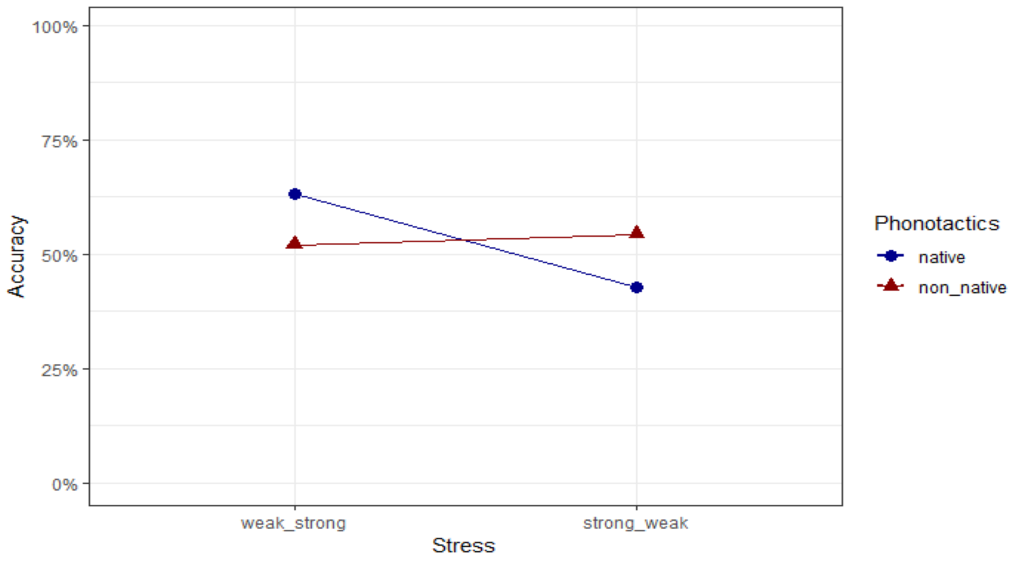

Figure 6-22. Forced-choice task: mean percentages of accuracy scores for interaction between phonotactics and stress conditions.

6.2.3.1.1 Linguistic group and the interaction of phonotactics with stress

For the word recognition task, a three-way ANOVA was conducted with Dprime as the dependent variable, and syllable stress and phonotactics as within-subject independent variables, and linguistic group as a between-subject independent variable so as to investigate whether or not the detection of words depends on the interaction between stress pattern, phonotactics, and linguistic group. Before the analysis, the data were aggregated according to Dprime score, subject, phonotactics, stress, and linguistic group variables. The main effect of this interaction was not significant [$F(1,24)=2.05$, $p=0.17$, $\eta_p^2=0.08$]. A pairwise comparison showed that, for the linguistically sophisticated group, weak-strong words with native phonotactics [$M=0.62$] were recognised significantly better than weak-strong words with non-native phonotactics [$M=-0.08$], [$p<0.01$]. There was also a statistically significant difference for the linguistically naïve group where weak-strong words with the native phonotactic pattern [$M=0.44$] were recognised better than weak-strong words with non-native phonotactics [$M=0.01$]. These results suggest that weak-strong words with the native phonotactic pattern were recognised better than weak-strong words with the non-native phonotactics for both groups of participants, but still the main effect of interaction between phonotactics, stress and linguistic group was not significant. The interaction of these variables is illustrated in Figure 6-23.

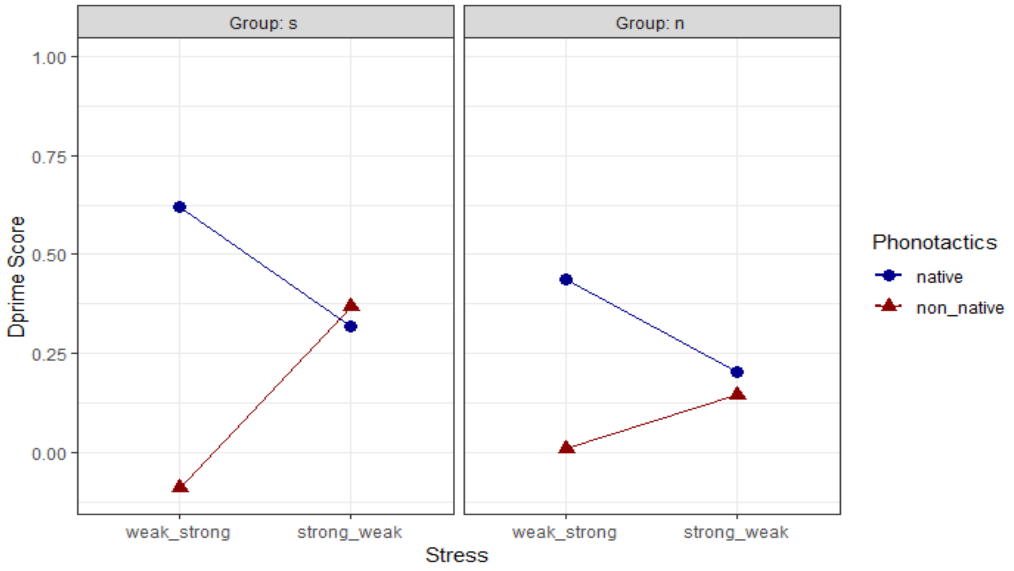

Figure 6-23. Word recognition task: Dprime scores for interaction between phonotactics and stress conditions by linguistic group: s=sophisticated group; n=naïve group.

Syllable Structure	Linguistically Sophisticated Group		Linguistically Naive Group	
	Phonotactics		Phonotactics	
	Native	Non-native	Native	Non-native
Weak-strong	0.62	-0.08	0.44	0.01
Strong-weak	0.32	0.37	0.20	0.15

Table 6-7. Word recognition task: Dprime scores for interaction between phonotactics and stress conditions by linguistic group: s=sophisticated group; n=naïve group.

Furthermore, to incorporate the effect of all levels of the phonotactics condition, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and phonotactics, word length and linguistic group as explanatory variables. The results of the model showed an insignificant effect of the interaction between naive group and words with non-native phonotactics and the strong-weak stress pattern [OR=0.85 (95% confidence interval [CI]: 0.60, 1.22), p=0.39], as well as the interaction between the naïve group, and words which followed phonotactics of neither language and the strong-weak stress pattern [OR=0.70 (95% confidence interval [CI]: 0.42, 1.15), p=0.16]. Figure 6-24 below illustrates this interaction, and the mean percentage accuracy values are presented in Table 6-8. Although the results are not statistically significant, however, figure 6-24 shows that there was no difference between the linguistic groups, and that the accuracy overall was higher for weak-strong words than strong-weak words when words had native phonotactics. Meanwhile, when words had non-native phonotactics, accuracy was higher for strong-weak words than weak-strong words. However, when phonotactics of words followed neither native or non-native, there was a difference between the linguistic group, such that the accuracy of the sophisticated group was higher for strong-weak rather than weak-strong words, but the accuracy of the naïve group was higher for weak-strong rather than strong-weak words.

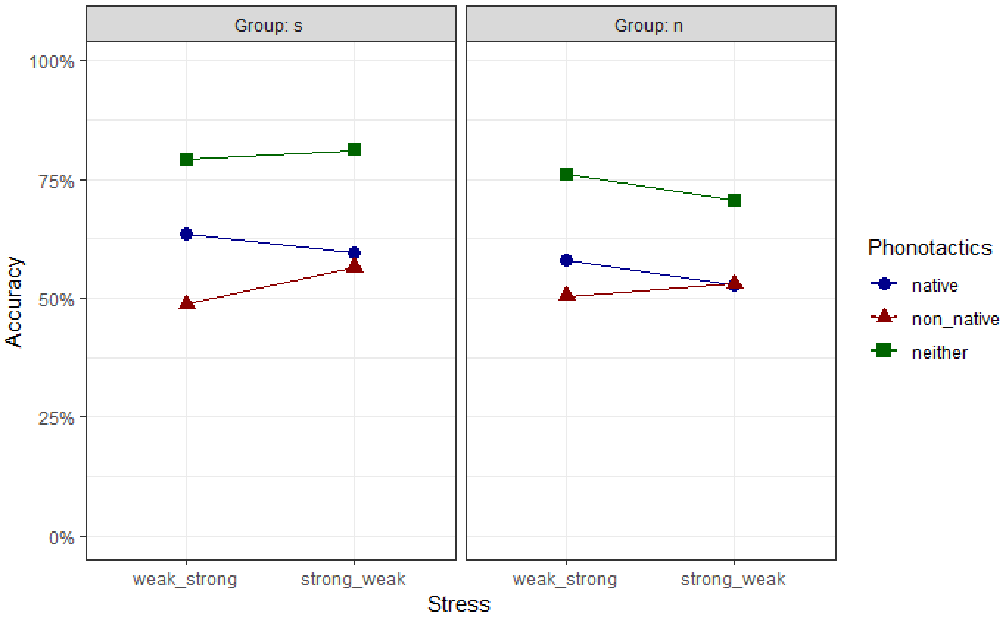

Figure 6-24. Word recognition task: mean percentages of accuracy scores for interaction between phonotactics and stress conditions.

Syllable Structure	Linguistically Sophisticated Group			Linguistically Naive Group		
	Phonotactics			Phonotactics		
	Native	Non-native	Neither	Native	Non-native	Neither
Weak-strong	62.9%	48.7%	78.2%	57.6%	50.3%	75.2%
Strong-weak	59%	56.2%	80.3%	52.5%	52.9%	69.7%

Table 6-8. Word recognition task: mean percentages of accuracy scores for interaction between phonotactics and stress conditions.

As for the forced-choice task, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and phonotactics, stress and linguistic group as explanatory variables in order to investigate if an ability to respond to novel words depends on the interaction between these variables.

The results of the model showed that the interaction between the naïve group, non-native phonotactics and syllable stress was not significant [*OR* =1.00 (95% confidence interval [*CI*]: 0.58, 1.74), *p*=0.98]. What this suggests is that, irrespective of language group, performance was better with weak-strong words than strong-weak words when the phonotactics of these words were native, whereas the performance of both groups was higher on strong-weak words than weak-strong words when the phonotactics were non-native Moreover, the performance of the sophisticated group was somewhat higher for weak-strong words with the native phonotactics [*M*=66%], compared to the linguistically naïve participants [*M*=59%]. Figure 6-25 below illustrates this interaction, and the values of mean percentage accuracy are presented in Table 6-9.

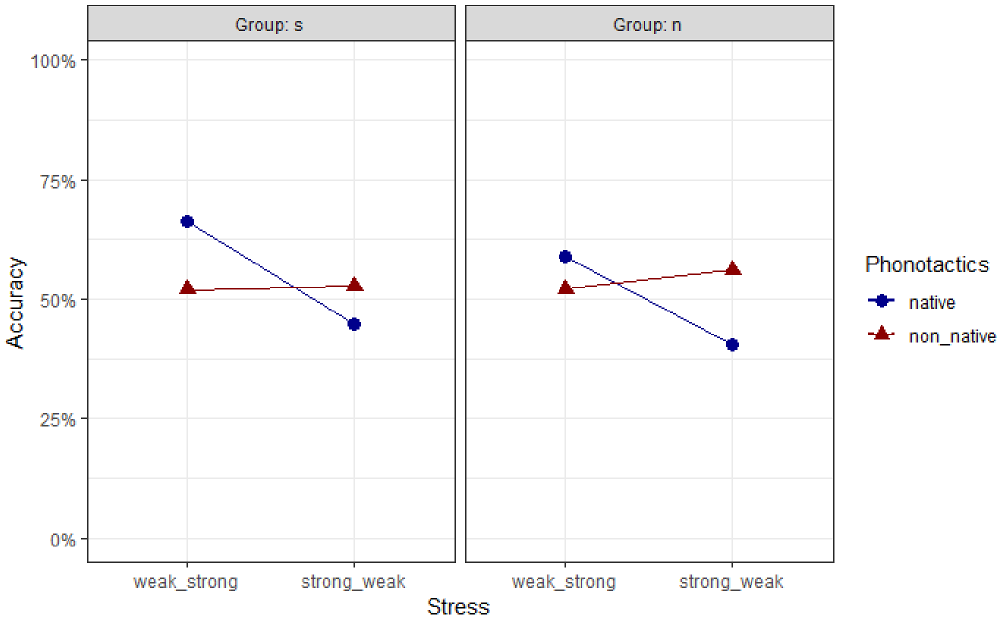

Figure 6-25. Forced-choice task: mean percentages of accuracy for interaction between phonotactics and stress conditions by linguistic group: s=sophisticated group; n=naïve group.

Syllable Structure	Linguistically Sophisticated Group		Linguistically Naive Group	
	Phonotactics		Phonotactics	
	Native	Non-native	Native	Non-native
Weak-strong	66%	52%	59%	52%
Strong-weak	45%	52%	41%	56%

Table 6-9. Forced-choice task: mean percentages of accuracy for interaction between phonotactics and stress conditions by linguistic group: s=sophisticated group; n=naïve group.

6.2.3.1.1.1 Summary

The results from the word recognition task and forced-choice tasks were significant. In particular, words which followed native phonotactics and were stressed on the final syllable were recognised better than words which followed native phonotactics and were of the strong-weak pattern. Moreover, the opposite appears to be true with words which followed non-native phonotactics, where strong-weak words were recognised better than weak-strong words. Words with non-native phonotactics and weak-strong stress pattern were recognised just about at a chance level. These results do not support *hypothesis 5,* which predicted that participants in this study would be more accurate in detecting Russian words from the input which have native phonotactics and which are stressed on the first syllable. The results showed that words with native phonotactics and weak-strong stress were the most accurately detected words across both tasks.

For neither task, there were significant differences between linguistic groups because the performance of both linguistically sophisticated and linguistically naïve participants was similar. In particular, both groups recognised words better if they followed native phonotactics and were stressed on the last syllable as opposed to strong-weak words with the native phonotactics. These results do not support *hypothesis 12,* which predicted that the accuracy of the linguistically sophisticated group would be higher than that of the linguistically naïve group in detecting strong-weak words with native phonotactics.

6.2.3.2 Interaction of phonotactics with length

For the word recognition task, a two-way ANOVA was conducted with Dprime as the dependent variable, and phonotactics (native vs non-native) and word length (monosyllabic vs bisyllabic) as independent variables so as to investigate whether or not an ability to respond to words from the input depends on the interaction of phonotactics and word length. Before the analysis, the data were aggregated according to Dprime score, subject, phonotactics and word length. The results revealed that this interaction was not significant [$F(1, 25) = 2.37$, p=0.14, η_p^2=0.09.]. Figure 6-26 below illustrates this interaction, and the means of Dprime scores for bisyllabic words are presented in Table 6-10. Despite statistically the insignificant results, the means of Dprime scores suggest that, when words had native phonotactics, the performance for bisyllabic words were slightly higher than for monosyllabic words, but when words had non-native phonotactics, performance was no affected by word length.

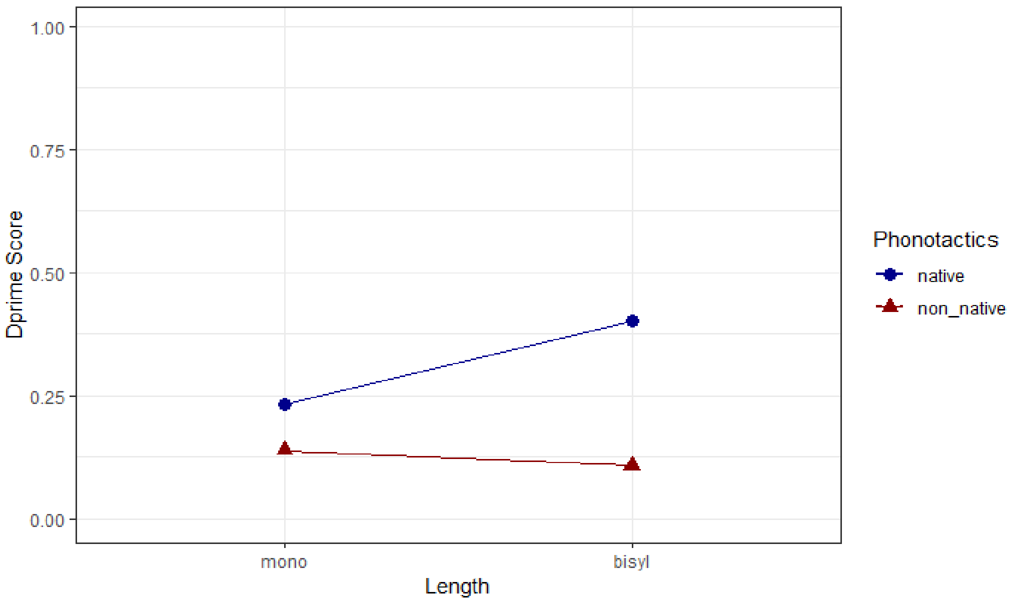

Figure 6-26. Word recognition task: Dprime scores for interaction between phonotactics and word length conditions: mono=monosyllabic words; bisyl=bisyllabic words.

Phonotactics	Word Length	
	Monosyllabic	Bisyllabic
Native	0.23	0.40
Non-native	0.13	0.11

Table 6-10. Word recognition task: Dprime scores on interaction between phonotactics and word length conditions: mono=monosyllabic words; bisyl=bisyllabic words.

Furthermore, to incorporate all levels of the phonotactics variable, a mixed-effect logistic regression model was fitted where the dependent variable was accuracy for a phonotactics condition, and the explanatory variables were all three levels of phonotactics and word length (monosyllables vs bisyllables). The results of the model showed that effect of any interaction between words with non-native phonotactics and monosyllabic words was not significant [OR=1 (95% confidence interval [CI]: 0.68, 1.45), p=0.99], and neither was that between words with phonotactics of neither language and monosyllabic words [OR=0.81 (95% confidence interval [CI]: 0.69, 1.45), p=0.40]. Figure 6-27 illustrates this interaction. The values of mean accuracy for this interaction are summarised in Table 6-11. The figure shows that accuracy for monosyllabic and bisyllabic words was nearly the same across words with native and non-native phonotactics. However, for the words with phonotactics of neither language, accuracy was better for bisyllabic [M=76%] rather than monosyllabic words [M=72.5%].

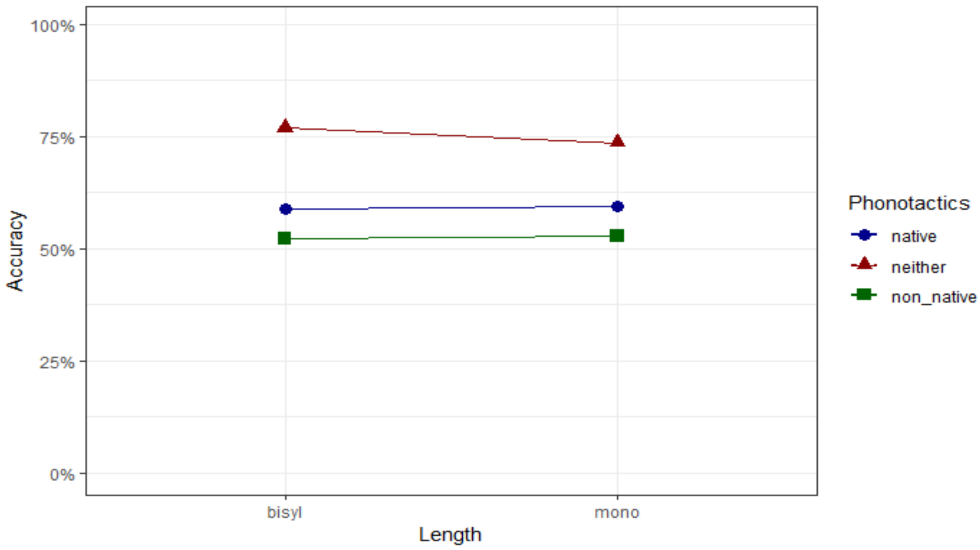

Figure 6-27. Word recognition task: mean percentages of accuracy scores on interaction between phonotactics and word length conditions.

Phonotactics	Word Length	
	Monosyllabic	Bisyllabic
Native	58.6%	58.1%
Non-native	52.6%	52.1%
Neither	72.5%	76.1%

Table 6-11. Word recognition task: mean percentages of accuracy scores for interaction between phonotactics and word length conditions.

For the forced-choice task, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable and word length and phonotactics as explanatory variables. The results showed that there was no statistically significant effect of an interaction between syllable length and phonotactics condition. The results for the interaction of the non-native phonotactics condition and the bisyllabic words were as follows: [*OR*=0.78 (95% confidence interval [*CI*]: 0.43, 1.45), *p*=0.44)]. Figure 6-28 below plots this interaction, and values of mean percentage accuracy are presented in Table 6-12.

Phonotactics	Word Length	
	Monosyllabic	Bisyllabic
Native	48.2%	52.8%
Non-native	54.1%	53%

Table 6-12. Forced-choice task: mean percentages of accuracy scores for interaction between phonotactics and word length conditions.

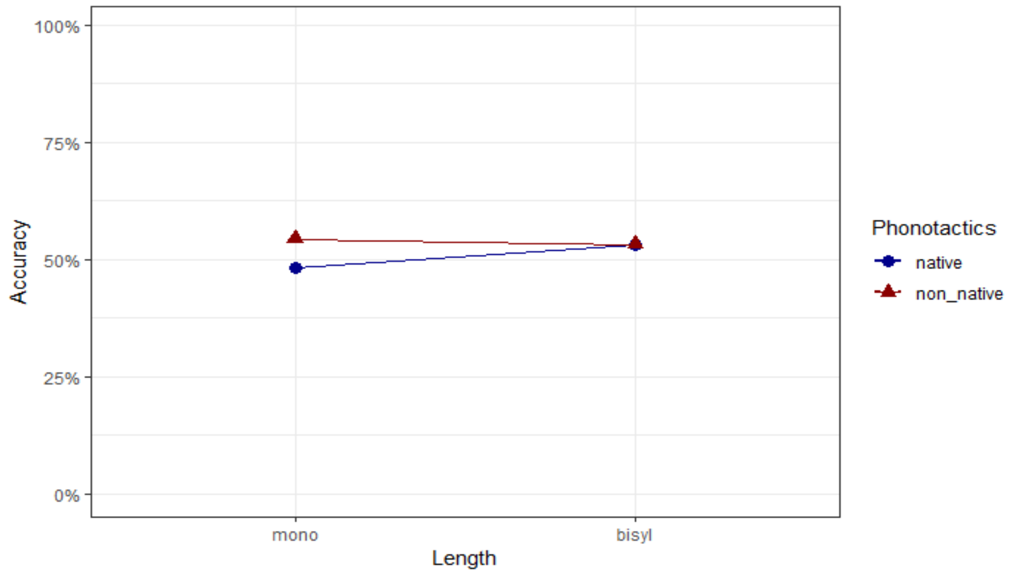

Figure 6-28. Forced-choice task: mean percentages of accuracy scores for interaction between phonotactics and word length conditions.

6.2.3.2.1 Linguistic group, and interaction of phonotactics with length

A three-way ANOVA was conducted with Dprime as the dependent variable, word length and phonotactics as within-subject independent variables and linguistic group as a between-subject independent variable in order to investigate if the detection of words depends on the interaction between syllable length, phonotactics, and linguistic group. Before the analysis, the data were aggregated according to Dprime score, subject, phonotactics, word length, and linguistic group variables. The analysis of variance showed that this interaction was not significant [$F(2,48)=1.79$, $p=0.18$, $\eta_p^2=0.07$]. Pairwise comparison showed that there were only two statistically significant differences for the linguistically sophisticated group, and in particular, monosyllabic words with native phonotactics [*M*=0.18] were recognised less often than bisyllabic words with native phonotactics [*M*=0.48]. Additionally, bisyllabic words with non-native phonotactics [*M*=0.14] were recognised less than bisyllabic words with native phonotactics [*M*=0.48]. Figure 6-29 below illustrates the results of this interaction and the mean Dprime scores are presented in Table 6-13.

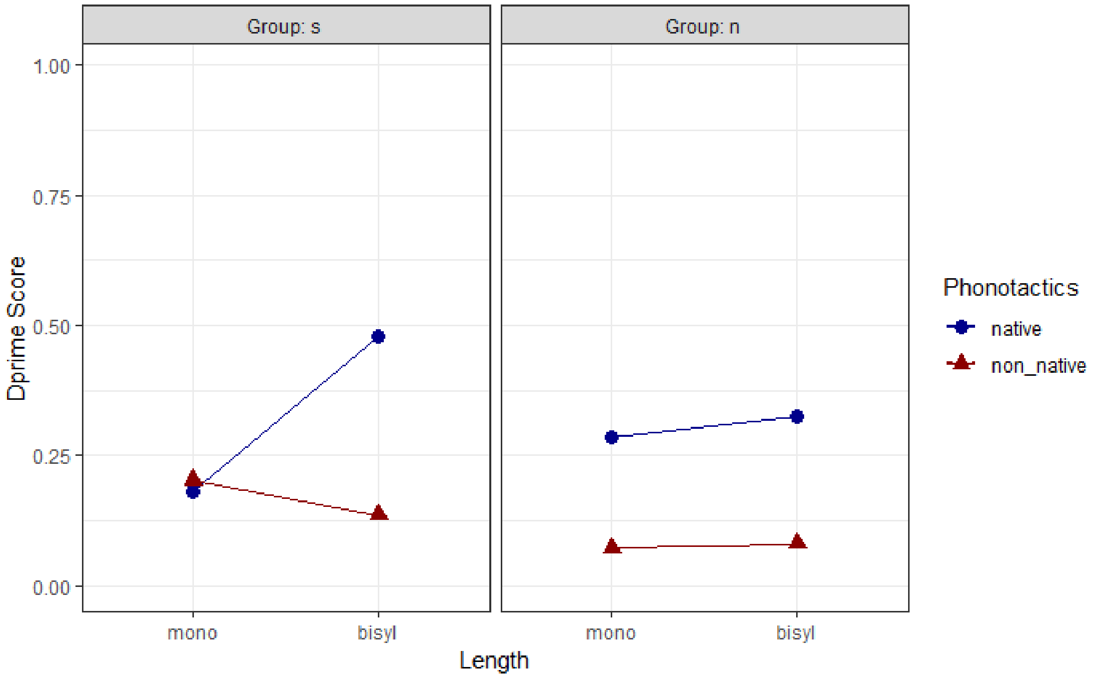

Figure 6-29. Word recognition task: Dprime scores for interaction between phonotactics and word length conditions by linguistic group: s=sophisticated group; n=naïve group.

Syllable Structure	Linguistically Sophisticated Group		Linguistically Naive Group	
	Phonotactics		Phonotactics	
	Native	Non-native	Native	Non-native
Monosyllables	0.18	0.20	0.28	0.07
Bisyllables	0.48	0.14	0.33	0.08

Table 6-13. Word recognition task: Dprime scores for interaction between phonotactics and word length conditions by linguistic group: s=sophisticated group; n=naïve group.

Furthermore, just as before, to incorporate the effect of all levels of the phonotactics condition, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and phonotactics, length and linguistic group as explanatory variables. The results of the model showed no significant effect of the interaction for the sophisticated group and monosyllabic words with neither English nor Russian phonotactics [OR=0.86 (95% confidence interval [CI]: 0.57, 1.33), *p*=0.51]. Moreover, there was a marginally significant effect of the interaction for the sophisticated group and monosyllabic words with non-native phonotactics [OR=1.31 (95% confidence interval [CI]: 0.96, 1.78), *p*=0.08]. Figure 6-30 below illustrates this interaction, and the mean percentage accuracy values are presented in Table 6-14. The figure shows that accuracy for monosyllabic and bisyllabic words was similar across words with native and non-native phonotactics for both linguistic groups, but accuracy in the sophisticated group was higher for bisyllabic words than for monosyllabic words when these words started with non-native phonotactics in the sophisticated group.

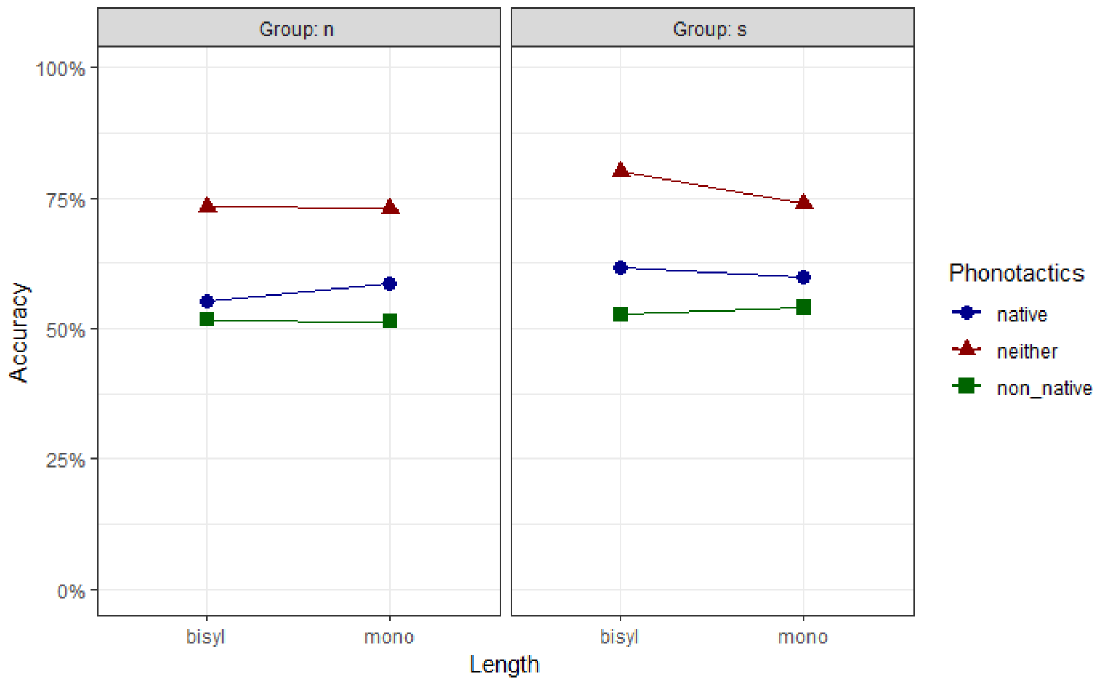

Figure 6-30. Word recognition task: mean percentages of accuracy scores for interaction between phonotactics and word length conditions by linguistic group: s=sophisticated group; n=naïve group; bisyl=bisyllabic words; mon=monosyllabic words.

Word Length	Linguistically Sophisticated Group			Linguistically Naive Group		
	Phonotactics			Phonotactics		
	Native	Non-native	Neither	Native	Non-native	Neither
Monosyllabic	59.1%	53.8%	72.9%	57.9%	52.5%	72.1%
Bisyllabic	60.9%	54.5%	79.2%	55%	51.6%	72.5%

Table 6-14. Word recognition task: mean percentages of accuracy scores for interaction between phonotactics and word length conditions by linguistic group.

For the forced-choice task, mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and phonotactics, word length and linguistic group as explanatory variables so as to investigate whether or not the ability to respond to novel words depends on an interaction between these variables. The results of the model showed that the interaction between naïve group, non-native phonotactics and bisyllabic words was not significant [OR=0.70 (95% confidence interval [CI]: 0.37, 1.35), *p*=0.32]. Figure 6-31 illustrates this interaction and the values of mean percentage accuracy are presented in Table 6-15.

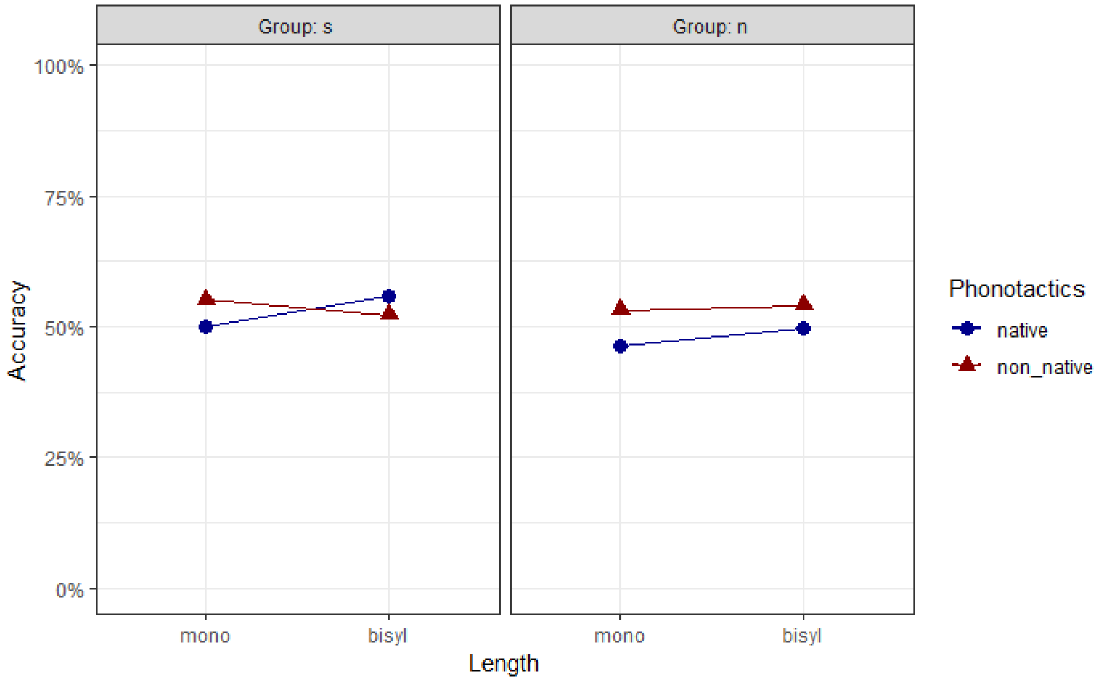

Figure 6-31. Forced choice task: mean percentages of accuracy scores for interaction between phonotactics and word length conditions by linguistic group: s=sophisticated group; n=naïve group; mono=monosyllabic words; bisyl=bisyllabic words.

Word Length	Linguistically Sophisticated Group		Linguistically Naive Group	
	Phonotactics		Phonotactics	
	Native	Non-native	Native	Non-native
Monosyllabic	50%	55%	46%	53%
Bisyllabic	55%	52%	50%	54%

Table 6-15. Forced choice task: mean percentages of accuracy scores for interaction between phonotactics and word length conditions by linguistic group.

6.2.3.2.1.1 Summary

None of the results from the word recognition and forced-choice tasks showed a significant interaction between phonotactics and word length. This finding does not support *hypothesis 6*, which predicted that participants would detect bisyllabic words better than monosyllabic words with native phonotactics.

Furthermore, there were no significant interactions among linguistic group, word length and phonotactics for either task. This once again does not support *hypothesis 12,* which predicted differences between the linguistically sophisticated and the linguistically naïve participants. The next section looks at whether or not sensitivity to phonotactics and stress and word length increases over sessions.

6.2.4 Interaction of cues over sessions

This section discusses the effects of the interactions of phonological cues over time. It starts by considering the interaction between phonotactics and session in Section 6.2.4.1, followed by the interaction between stress and session in Section 6.2.4.2. Section 6.2.4.3 then looks at the interaction between word length and session. Each section is followed by a discussion of the effect of linguistic group on each of these interactions.

6.2.4.1 Phonotactics and session

A two-way ANOVA was conducted with Dprime as the dependent variable and phonotactics condition (native vs non-native) and session as independent variables so as to investigate whether or not sensitivity to phonotactic constraints of words from the input increases over sessions. Before the analysis, the data were aggregated according to Dprime score, subject, session and phonotactics.

The analysis of variance revealed that the interaction of phonotactics and session was not significant [$F(3, 75) = 1.23$, $p=0.30$, $\eta_p^2=0.05$]. Although the *p*-value is far from being significant, Figure 6-32 below suggests that the interaction between phonotactics and session did slightly influence the participants' Dprime scores. That is why a posthoc analysis of by phonotactics condition was subsequently conducted. A pairwise comparison showed, that for native phonotactics, there were statistically significant differences between sessions 1 and 3 [$p<0.01$], and between sessions 1 and 4 [$p<0.01$], as well as between sessions 2 and 4 [$p<0.01$], but there were no statistically significant differences between sessions 1 and 2 [$p=0.12$], or between sessions 2 and 3 [$p=0.12$], while the difference between sessions 3 and 4 was marginally significant [$p=0.08$]. Meanwhile, pairwise comparisons for the non-native phonotactics condition showed that there were only two statistically significant differences, which were between sessions 1 and 2 [$p=0.049$], and between sessions 1 and 4 [$p<0.01$]. In other words, there were no statistically significant differences between sessions 1 and 3 [$p=0.13$], between sessions 2 and 3 [$p=0.64$], while the difference between sessions 3 and 4 was marginally significant [$p=0.08$]. Table 6-16 below presents the mean *d'* scores for each interaction between phonotactics and session. Although no statistically significant main effect of an interaction between session and phonotactics was found, Figure 6-32 and the mean Dprime scores illustrate that there is an overall trend of increasing scores for both native and non-native phonotactics, and that the mean Dprime scores are higher for words with the native phonotactics than non-native ones across all sessions.

Phonotactics	Session			
	1	2	3	4
Native	0.14	0.28	0.41	0.57
Non-native	-0.04	0.14	0.10	0.25

Table 6-16. Word recognition task: Dprime scores for interaction between session and phonotactics condition.

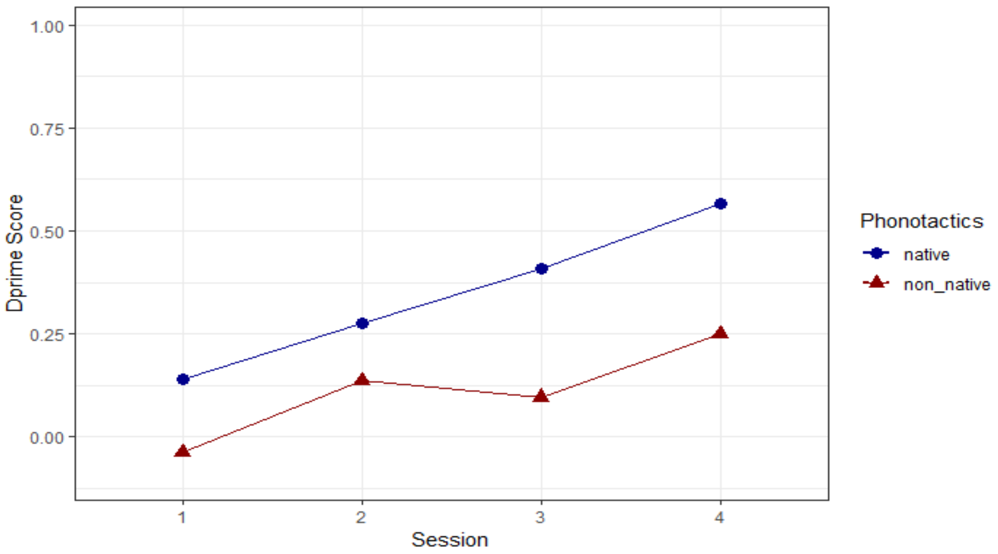

Figure 6-32. Word recognition task: Dprime scores for interaction between session and phonotactics condition.

Furthermore, to gain a clearer understanding of the role of all levels of the phonotactics variable, a mixed-effect logistic regression model was fitted for the dependent variable of accuracy, and explanatory variables of phonotactics condition with all three levels and session. The results are summarised in Table 6-17 below, indicating a statistically insignificant interaction between these variables. The mean scores for this interaction are presented in Table 6-18 below, and Figure 6-33 illustrates them. Although no single interaction between phonotactics and session has a significant effect, it is evident from the means that the effect of phonotactics on accuracy increases with more input for all levels of the phonotactics condition. However, there is a decrease in accuracy scores for all levels of the phonotactics variable in session 3. The highest accuracy among all levels of the phonotactics condition is when the words followed the phonotactics of neither English nor Russian. This is because the participants were correctly rejecting these items which represented non-generalisable distractors. The next highest accuracy was for words with native phonotactics, and then words with non-native phonotactics, which received the lowest accuracy scores. The trend for non-native and native phonotactics is the same as in the above discussion of the ANOVA analysis.

Interaction of variables	OR (Odds Ratios)	CI (Confidence Intervals)	*p*-value
session2:Phonotacticsneither	1.32	0.99; 1.75	0.05
session3:Phonotacticsneither	0.97	0.73; 1.28	0.82
session4:Phonotacticsneither	1.07	0.80; 1.41	0.65
session2:Phonotacticsnon_native	1.01	0.82; 1.24	0.90
session3:Phonotacticsnon_native	0.88	0.72; 1.08	0.22
session4:Phonotacticsnon_native	0.86	0.70; 1.04	0.13

Table 6-17. Word recognition task: results from mixed-effect logistic regression for interaction between session and phonotactics condition.

Phonotactics	Session			
	1	2	3	4
Native	54.3%	56.9%	59.4%	62.5%
Non-native	49.7%	52.7%	52%	54.6%
Neither	70%	77.3%	73.9%	78%

Table 6-18. Word recognition task: mean percentages of accuracy scores for interaction between session and phonotactics condition.

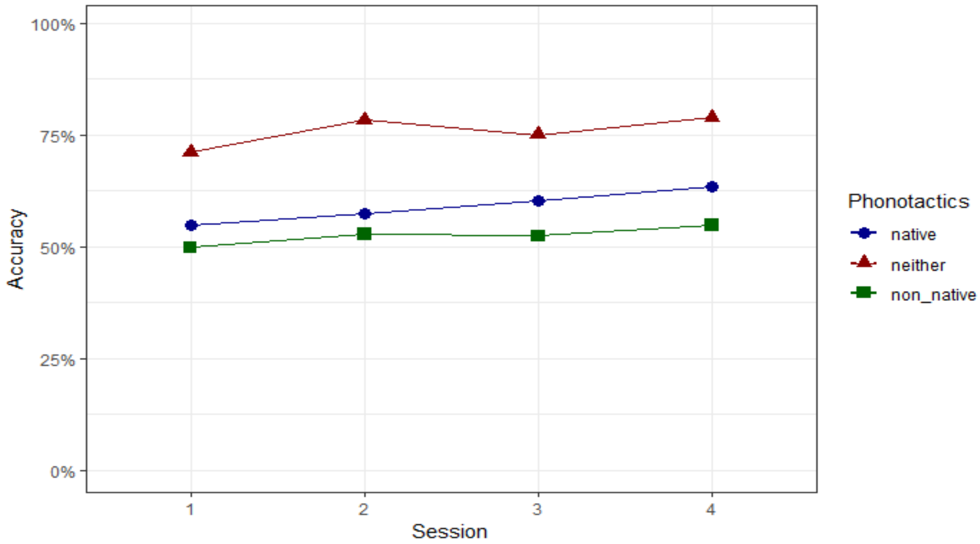

Figure 6-33. Word recognition task: mean percentages of accuracy scores on interaction between session and phonotactic condition.

Now, turning to the results from the forced-choice task, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and session and phonotactics condition as explanatory variables in order to investigate whether or not sensitivity to phonotactic constraints increases over sessions. The results for this model showed no significant effect of an interaction between session 2 and non-native phonotactics [OR=1.07 (95% confidence interval [CI]: 0.78, 1.47), p=0.65], which means that accuracy was only slightly better on the session 2 than session 1 for the non-native phonotactics. Furthermore, there was no significant effect for session 3 for non-native phonotactics [OR =0.98 (95% confidence interval [CI]: 0.72, 1.34), p=0.91]. Finally, the effect was also not significant between session 4 and non-native phonotactics [OR =0.95 (95% confidence interval [CI]: 0.70, 1.30), p=0.76]. The mean percentage accuracy values for the phonotactics condition per each session are demonstrated in Table 6-19, and Figure 6-34 illustrates the interaction between these variables.

Session	1	2	3	4
Phonotactics condition=native	48%	49.5%	51.3%	56.2%
Phonotactics condition=non-native	50%	53.2%	52.9%	57.2%

Table 6-19. Forced-choice task: mean percentages of accuracy scores for interaction between session and phonotactics condition.

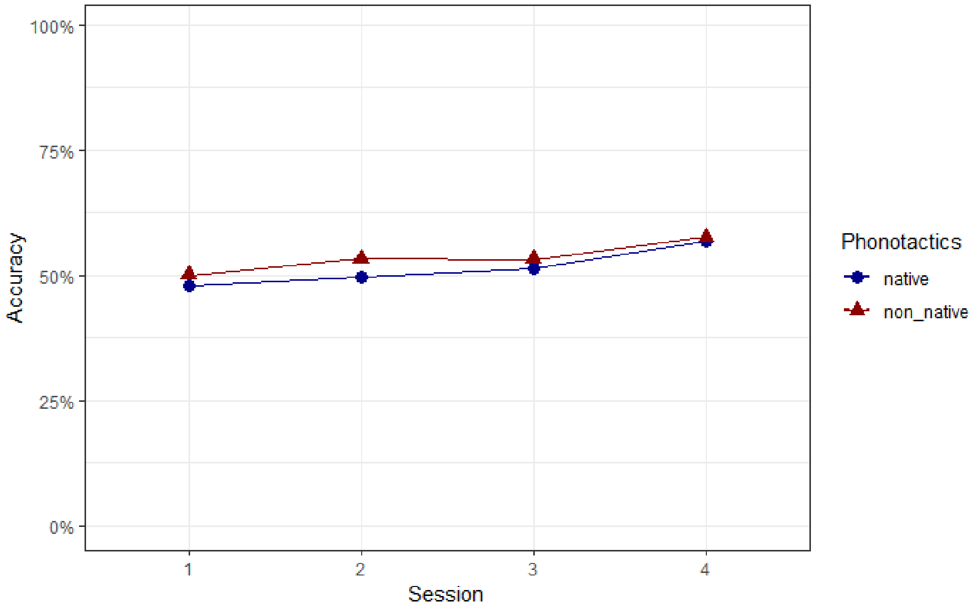

Figure 6-34. Forced-choice task: mean percentages of accuracy scores for interaction between session phonotactics condition.

6.2.4.1.1 Linguistic group, phonotactics and session

A three-way ANOVA was conducted with Dprime as the dependent variable and phonotactic nativeness condition (native vs non-native) and session (with all four levels) as within-subject independent variables and linguistic group as a between-subject independent variable in order to investigate whether or not sensitivity to phonotactic constraints of words from the input increases over sessions, and whether there is a difference related to linguistic group. Before the analysis, the data were aggregated according to Dprime score, subject, session, phonotactics, and linguistic group variables.

The results showed that this interaction was not significant [$F(3,72)=0.21$, $p=0.89$, $\eta_p^2=0.009$]. Figure 6-35 below illustrates the interaction between phonotactics condition, session and linguistic group. The figure shows that, although the linguistically sophisticated group performed slightly better on words with native phonotactics than the linguistically naïve group, both groups showed a similar pattern of development where their detection of words with native phonotactics improved on each day. For words with native phonotactics, a pairwise comparison showed a significant difference between sessions 1 [$M=0.10$] and 4 [$M=0.62$] for the sophisticated group [$p<0.01$], whereas there was no significant difference for the linguistically naïve group between sessions 1 [$M= -0.04$] and 4 [$M=0.52$] [$p=0.13$]. Additionally, for the linguistically sophisticated group, words with non-native phonotactics in sessions 1 [$M=-0.04$], 2 [$M=0.14$], and 3 [$M=0.18$] were recognised significantly less often than words with native phonotactics in session 4 [$M=0.61$], with all p-values being <0.05. These results suggest that words with native phonotactics were recognised best of all by the linguistically sophisticated group. The mean Dprime scores for the phonotactics condition across the four sessions is presented in Table 6-20 for each group.

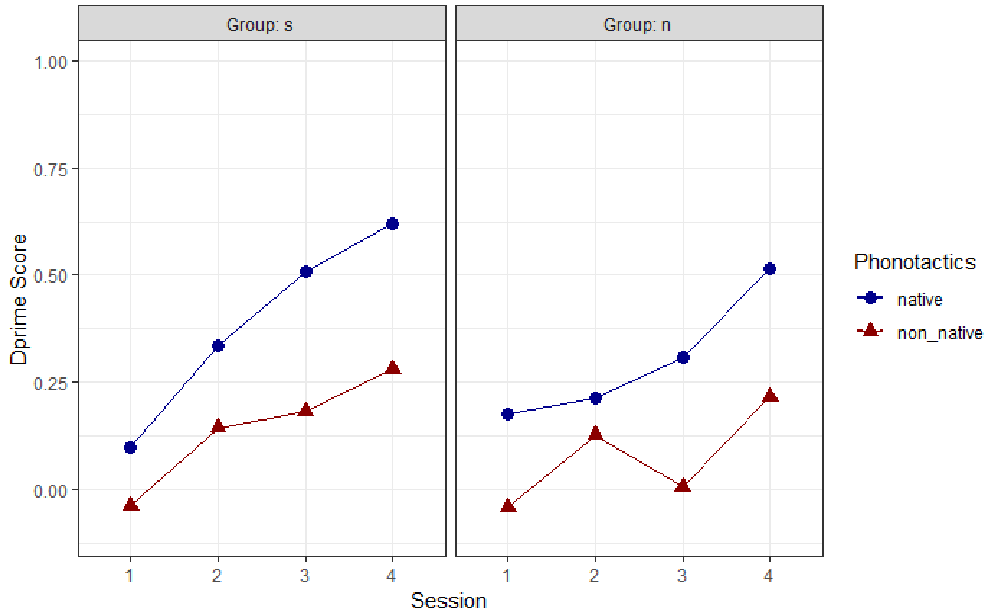

Figure 6-35. Word recognition task: Dprime scores for interaction between session, phonotactics condition by linguistic group: s=sophisticated group; n-naïve group.

Session	Linguistically Sophisticated Group		Linguistically Naive Group	
	Phonotactics		Phonotactics	
	Native	Non-native	Native	Non-native
1	0.10	-0.04	0.18	-0.04
2	0.34	0.15	0.21	0.13
3	0.51	0.18	0.31	0.01
4	0.62	0.28	0.52	0.22

Table 6-20. Word recognition task: Dprime scores for interaction between session and phonotactics condition by linguistic group: s=sophisticated group; n-naïve group.

Additionally, a mixed-effect logistic regression model was fitted were the dependent variable was accuracy, and explanatory variables were the phonotactics condition with all three levels (native, non-native, and neither) and session. The results for the model are shown in Table 6-21, and figure 6-36 below illustrates this interaction with the mean percentage accuracy presented in Table 6-22. The results confirm those from the analysis of Dprime scores in ANOVA in that there is no statistically significant difference between the sophisticated and the naïve group. However, there was a trend where both groups showed improvement and recognised words with native phonotactics better than words with non-native phonotactics from session 1 to session 4, with the linguistically sophisticated group showing a tendency for higher accuracy in word detection.

Variables interaction	*OR* (Odds Ratios)	*CI* (Confidence Intervals)	*p*-value
session2:Phonotacticsneither:Groups	0.89	0.51; 1.57	0.70
session3:Phonotacticsneither:Groups	0.66	0.38; 1.17	0.16
session4:Phonotacticsneither:Groups	1.07	0.61; 1.88	0.82
session2:Phonotacticsnon_native:Groups	0.79	0.53; 1.19	0.26
session3:Phonotacticsnon_native:Groups	0.86	0.57; 1.30	0.49
session4:Phonotacticsnon_native:Groups	0.79	0.53; 1.19	0.26

Table 6-21. Word recognition task: results from mixed-effect logistic regression for interaction between phonotactics condition and session by linguistic group: Groups=sophisticated group.

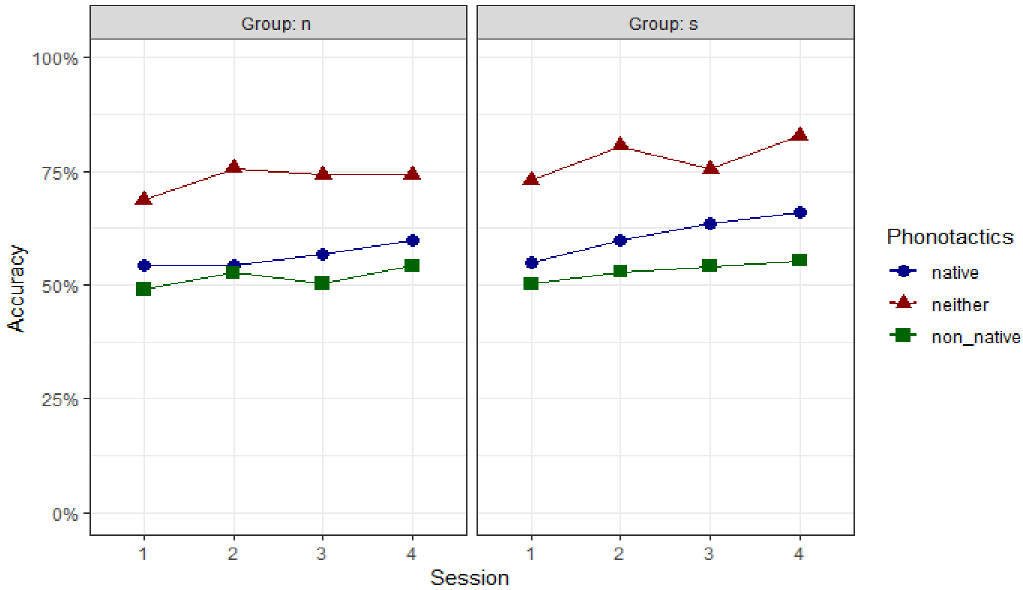

Figure 6-36. Word recognition task: mean percentages of accuracy scores for interaction between phonotactic condition and session by linguistic group: s=sophisticated group; n=naïve group.

Session	Linguistically Sophisticated Group			Linguistically Naive Group		
	Phonotactics			Phonotactics		
	Native	Non-native	Neither	Native	Non-native	Neither
1	54%	49.2%	67.9%	54.5%	50%	72%
2	54.2%	52.6%	74.7%	59.2%	52.8%	79.7%
3	56.4%	50.3%	73.4%	62.5%	53.7%	74.4%
4	59.4%	54%	73.4%	65.2%	55.1%	81.9%

Table 6-22. Word recognition task: mean percentages of accuracy scores for interaction between phonotactics condition and session by linguistic group.

For the forced-choice task, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and session, phonotactics and linguistic group as explanatory variables to investigate whether or not sensitivity to phonotactic constraints of words from the input increases over sessions and whether there is an effect of linguistic group. The model produced the following results: (1) for the interaction between the naïve group, the non-native phonotactics and session 2 [*OR* =1.59 (95% confidence interval [*CI*]: 0.84, 2.99), $p<0.15$]; (2) for the interaction between naïve group, non-native phonotactics and session 3 [*OR* =1.50 (95% confidence interval [*CI*]: 0.80, 2.80), $p<0.20$]; (3) for the interaction between naïve group, non-native phonotactics and session 4 [*OR* =0.79 (95% confidence interval [*CI*]: 0.42, 1.48), $p<0.47$]. Mean percentage accuracy values for this interaction are presented in Table 6-23. Figure 6-37 indicates that, for the linguistically sophisticated group, words with native phonotactics were recognised slightly better than words with non-native phonotactics throughout the sessions, but on session 4 participants scored slightly higher with non-native phonotactics. On the other hand, the linguistically naïve participants' performance was more accurate for non-native phonotactics until session 3, and by session 4 the levels of recognition of words with native and non-native phonotactics were similar.

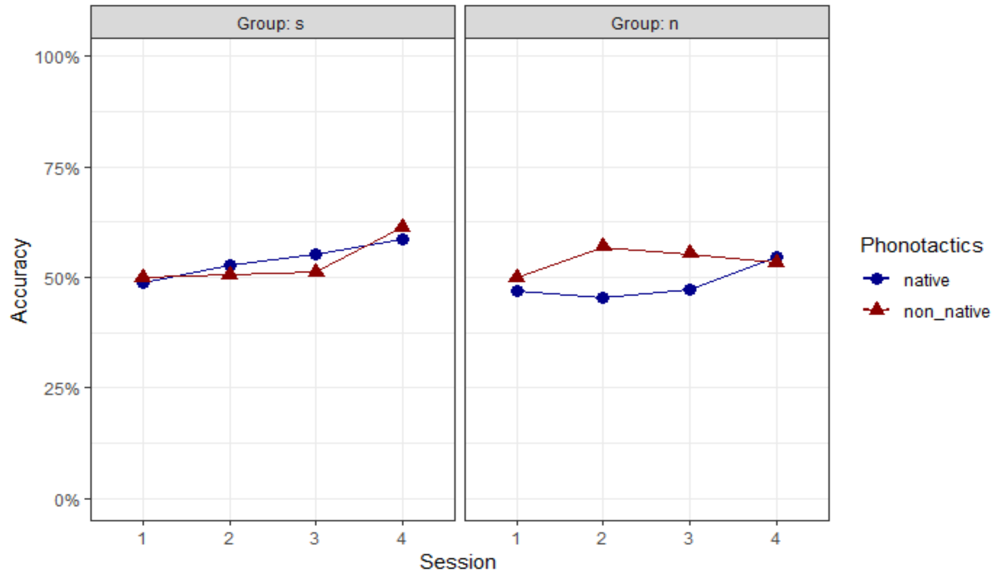

Figure 6-37. Forced-choice task: mean percentages of accuracy scores for interaction between session and phonotactics condition by linguistic group: s=sophisticated group; n=naïve group.

Session	Linguistically Sophisticated Group		Linguistically Naive Group	
	Phonotactics		Phonotactics	
	Native	Non-native	Native	Non-native
1	48.8%	50%	47.1%	50%
2	52.5%	50.5%	45.8%	56.5%
3	54.7%	51.1%	47.4%	55.1%
4	57.9%	60.8%	54.1%	53.2%

Table 6-23. Forced-choice task: mean percentages of accuracy scores on interaction between session and phonotactic condition by linguistic group. s=sophisticated group; n=naïve group.

6.2.4.1.1.1 Summary

To sum up, in neither the word recognition nor the forced-choice tasks were any significant interactions found between the levels of the session variable and these of the phonotactic condition variable. However, in both tasks it was found that the word detection ability for both native and non-native phonotactics conditions improved with more input, which is similar to the discussion concerning the effect of sessions on the participants' ability to detect the words. Furthermore, the results for the word recognition task showed that words with native phonotactics tended to be detected more successfully than words with non-native phonotactics, while the results for the forced-choice task showed the opposite pattern, where words with non-native phonotactics were detected better than words with native phonotactics. All in all, given that the overall differences were not statistically significant, *hypothesis 7,* which predicted that there would be an interaction between session and phonotactics is not supported for either task.

There was also no statistically significant difference to indicate that the performance of the sophisticated group with words with native and non-native phonotactics differed significantly from that of naïve participants with more input. However, there was a trend where both groups showed improvements, and recognised words with native phonotactics better than words with non-native phonotactics from session 1 to session 4 in the word recognition task, whereas the was no such clear trend in the forced-choice task, except for an increase in accuracy scores across both levels of the phonotactics condition. These results, once again, do not support *hypothesis 12*.

6.2.4.2 Stress and session

A two-way ANOVA was conducted with Dprime as the dependent variable, and session and stress (weak-strong vs strong-weak) as independent variables to investigate whether or not an ability to respond to words from the input depends on an interaction between stress and session. Before the analysis, the data were aggregated according to Dprime score, subject and session and stress. The analysis of variance revealed that the main effect of interaction between stress and session was not significant [$F(3, 75) = 1.29$, $p=0.28$, $\eta_p^2=0.05$]. Table 6-24 below presents the mean *d'* scores for each interaction, and Figure 6-38 illustrates this interaction. It is evident that, despite the insignificant main interaction of session and stress, the mean Dprime scores increased across all sessions for both stress patterns. However, the mean Dprime scores are higher for weak-strong words than strong-weak words across all sessions except for the first day where the pattern is reversed.

Stress	Session			
	1	2	3	4
Strong-weak	0.13	0.18	0.28	0.38
Weak-strong	0.06	0.26	0.30	0.53

Table 6-24. Word recognition task: Dprime scores for interaction between session and stress conditions.

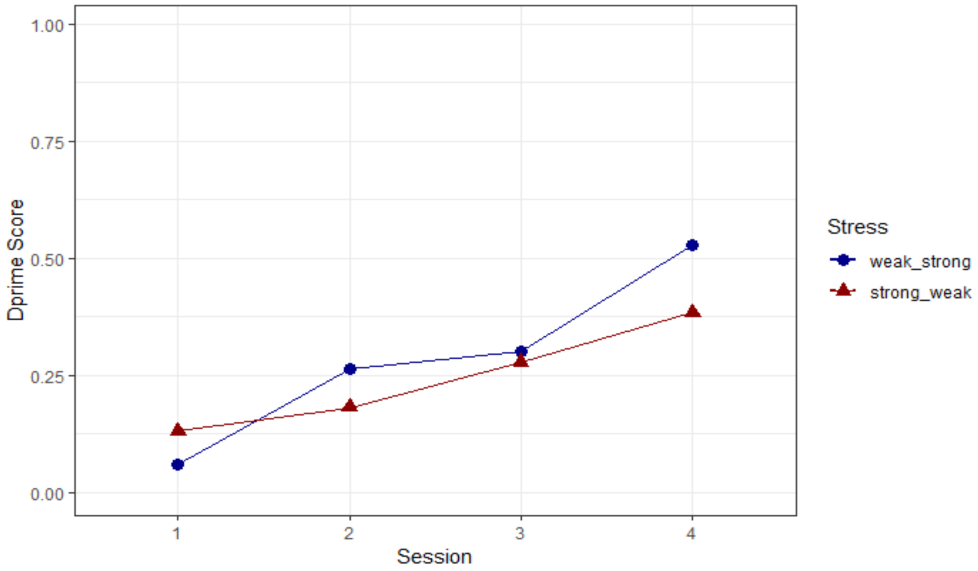

Figure 6-38. Word recognition task: Dprime scores for interaction between session and stress condition.

Finally, another model was fitted to investigate the interaction between stress and session. The results of this model showed that there was no significant interaction between any levels of these explanatory variables: (1) [OR = 0.92 (95% confidence interval [CI]: 0.62, 1.35), *p*=0.65] for session 2 and words with the strong-weak pattern; (2) [OR = 0.80 (95% confidence interval [CI]: 0.55, 1.18), *p*=0.27] for session 3 and strong-weak words; (3) and also [OR = 0.90 (95% confidence interval [CI]: 0.61, 1.32), *p*=0.59] for session 4 and strong-weak words. Although the results were not significant, the OR values less than one suggest that accuracy for words with the weak-strong stress pattern was higher than for words with the weak-strong stress pattern across the sessions. Table 6-25 below summarises the mean percentage accuracy against stress condition across the four sessions, and Figure 6-39 illustrates this interaction.

Stress	Session			
	1	2	3	4
Strong-weak	46.8%	47.4%	46.8%	53%
Weak-strong	53.1%	55.7%	58.3%	61.7%

Table 6-25. Forced-choice task: accuracy scores for interaction between session and stress condition.

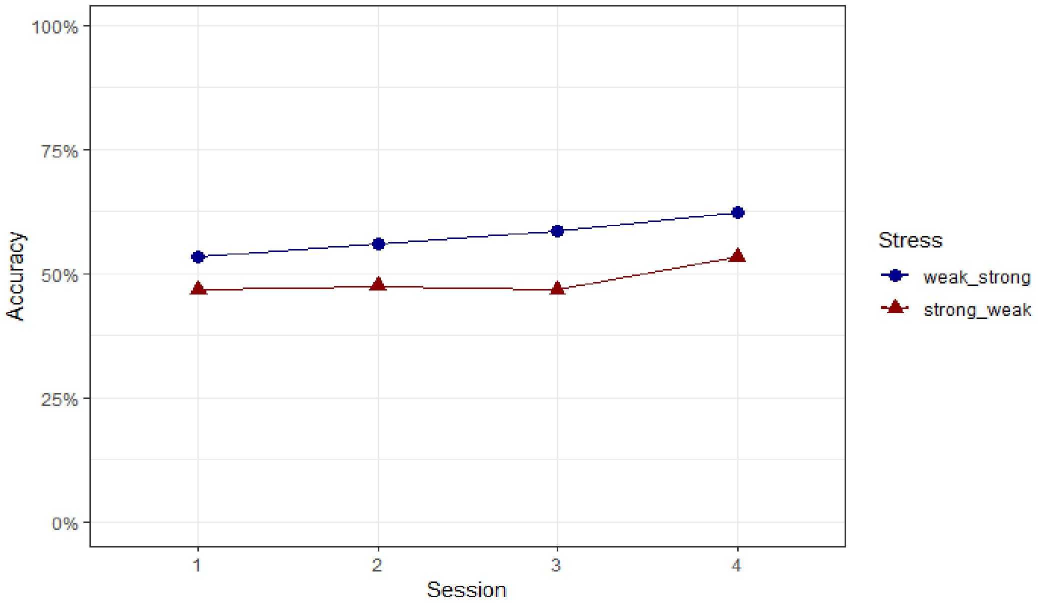

Figure 6-39. Forced-choice task: mean percentages of accuracy scores for interaction between session and stress condition.

6.2.4.2.1 Linguistic group, stress and session

A three-way ANOVA was conducted with Dprime as the dependent variable, syllable stress, and session as within-subject independent variables, and linguistic group as a between-subjects independent variable so as to investigate whether or not an ability to detect new words depends on the interaction of stress, session, and linguistic group. Before the analysis, the data were aggregated according to Dprime score, subject, session, stress, and linguistic group variables. The results of the analysis of variance for the main effect showed that this interaction was not significant [F(3,72)=1.73, p=0.17, η_p^2=0.07]. Pairwise comparison showed that, for the linguistically sophisticated group, there was a significant difference between performance on weak-strong words in sessions 1 [M=0.03] and 4 [M=0.54] with p<0.01, whereas for the linguistically naive group the difference between sessions 1 [*M*=0.09] and 4 [*M*=0.51] was only marginally significant [*p*=0.07]. No other meaningful pairwise comparison results were observed. These results indicate that, although the main effect of the interaction among stress, session and linguistic group was not significant, there was a trend towards the better detection of weak-strong words irrespective of whether participants were in the linguistically sophisticated or linguistically naïve group. Figure 6-40 below illustrates this interaction, and the mean Dprime scores are presented in Table 6-26.

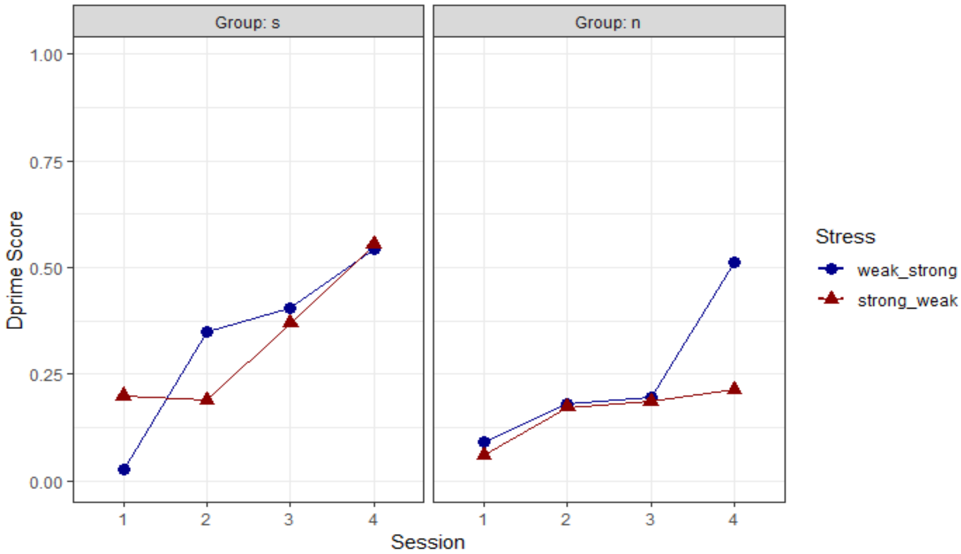

Figure 6-40. Word recognition task: Dprime scores for interaction between session and stress condition by linguistic group: s=sophisticated group; n=naïve group.

Session	Linguistically Sophisticated Group		Linguistically Naive Group	
	Stress		Stress	
	Weak-strong	Strong-weak	Weak-strong	Strong-weak
1	0.03	0.20	0.09	0.06
2	0.35	0.19	0.18	0.17
3	0.40	0.37	0.20	0.19
4	0.54	0.55	0.51	0.21

Table 6-26. Word recognition task: Dprime scores for interaction between session and stress condition by linguistic group: s=sophisticated group; n=naïve group.

Finally, for the forced-choice task, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and session, stress condition and linguistic group as explanatory variables. This model's interactions were not significant with the following results: (1) for the interaction between the naïve group, session 2, and the strong-weak pattern [OR =1.05 (95% confidence interval [CI]: 0.48, 2.88), p=0.90]; (2) for the interaction between the naïve group, session 3, and the strong-weak pattern [OR = 0.73 (95% confidence interval [CI]: 0.34, 1.57), p=0.42]; (3) and finally for the interaction between the naïve group session 4 and the strong-weak pattern [OR = 1.40 (95% confidence interval [CI]: 0.57, 2.70), p=0.58]. Figure 6-41 illustrates this interaction. Despite the statistically insignificant results, both groups recognised weak-strong words more accurately than strong-weak words, and there was a general increase in accuracy scores from sessions 1 to 4. Moreover, the performance of the linguistically sophisticated group by session 4 was better than the performance of the naïve group. Mean percentage accuracy scores for the interaction between session, stress condition and linguistic group are presented in Table 6-27.

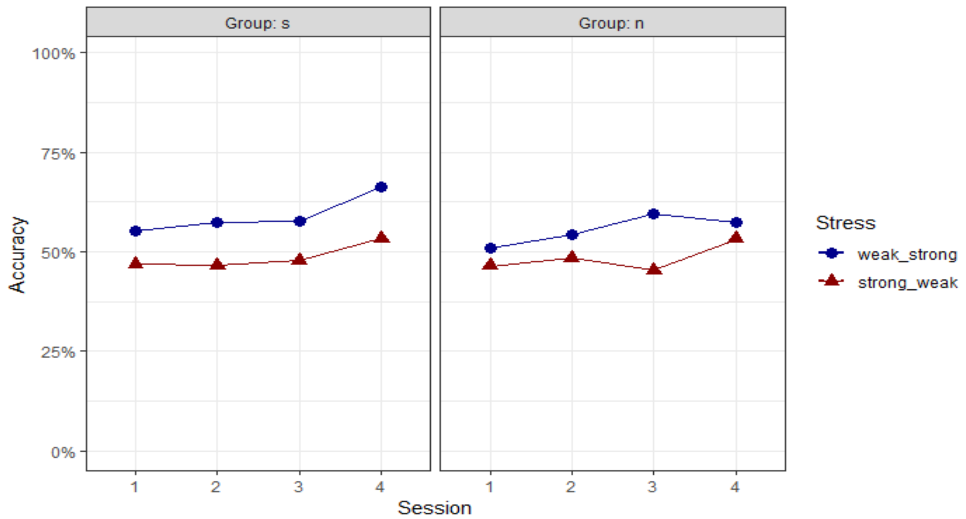

Figure 6-41. Forced-choice task: mean percentages of accuracy scores for interaction between session and stress condition by linguistic group: s=sophisticated group; n=naïve group.

Session	Linguistically Sophisticated Group		Linguistically Naive Group	
	Stress		Stress	
	Weak-strong	Strong-weak	Weak-strong	Strong-weak
1	55%	47%	51%	46.6%
2	57%	46.6%	54.1%	48.4%
3	57.5%	47.9%	59.1%	45.6%
4	65.7%	53.1%	57.2%	52.8%

Table 6-27. Forced-choice task: mean percentages of accuracy scores for interaction between session and stress condition by linguistic group: s=sophisticated group; n=naïve group.

6.2.4.2.1.1 Summary

To sum up, the results for the word recognition task and forced-choice tasks showed that there were no statistically significant effects of the interaction of session and stress variables. However, these interactions exhibited similar tendencies, in the sense that participants' accuracy was higher in detecting weak-strong words rather than strong-weak words. Additionally, there was a common trend of improvement trend from session 1 to session 4 with an occasional drop in accuracy in session 3, which subsequently rose again. These results do not support either *hypothesis 8*, which predicted that participants would be more accurate in detecting words which are stressed on the first syllable than on the second syllable and that this ability would increase over sessions.

The results for both tasks found no statistically significant effect of linguistic group. However, the results of the analysis showed that there was a tendency for weak-strong words to be recognised better than strong-weak words, and the recognition of words improved throughout the sessions for both linguistically sophisticated and naïve participants. These results again do not support *hypothesis 12*.

6.2.4.3 Length and session

For the word recognition task, a two-way ANOVA was conducted with Dprime score as the dependent variable, and session and word length (monosyllabic vs bisyllabic) as independent variables in order to investigate whether or not an ability to respond to words from the input depends on an interaction between word length and session. Before the analysis, the data were aggregated according to Dprime score, subject and session and word length variables. The results showed no significant interaction as the main effect between word length and session [$F(3, 75) = 0.11$, $p=0.951$, $\eta_p^2=0.005$]. Table 6-28 below provides the mean Dprime scores for word length across all sessions. Figure 6-42 shows that, despite statistically insignificant results, bisyllabic words were detected slightly more often over monosyllabic words across all four sessions.

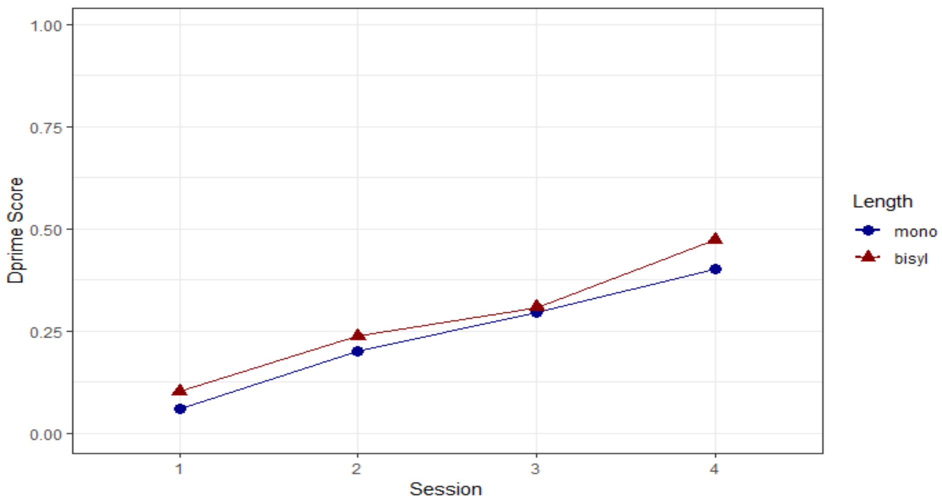

Figure 6-42. Word recognition task: Dprime scores for interaction between session and word length condition: mono=monosyllabic words; bisyl=bisyllabic words.

Word Length	Session			
	1	2	3	4
Monosyllables	0.06	0.20	0.30	0.40
Bisyllables	0.10	0.24	0.31	0.47

Table 6-28. Word recognition task: Dprime scores for interaction between session and word length condition: mono=monosyllabic words; bisyl=bisyllabic words.

For the forced-choice task, a mixed-effect logistic regression model was fitted to investigate whether or not there was any effect of interaction between word length and session on participants ability to detect words from the input. The results showed again that there was no significant interaction between the explanatory variables: (1) for session 2 and bisyllabic words [OR = 0.92

(95% confidence interval [CI]: 0.66, 1.27), *p*=0.60]; (2) for session 3 and bisyllabic words [OR = 0.94 (95% confidence interval [CI]: 0.67, 1.30), *p*=0.70]; (3) for session 4 and bisyllabic words and also [*OR* = 0.97 (95% confidence interval [*CI*]: 0.70, 1.35), *p*=0.86]. Figure 6-43 below illustrates this interaction. The mean values of percentage accuracy are presented in Table 6-29. These results suggest that there was no effect of an interaction between session and word length, but there was a general trend of improvement for both monosyllabic and bisyllabic words. However, there was a trend for bisyllabic words to be recognised slightly better.

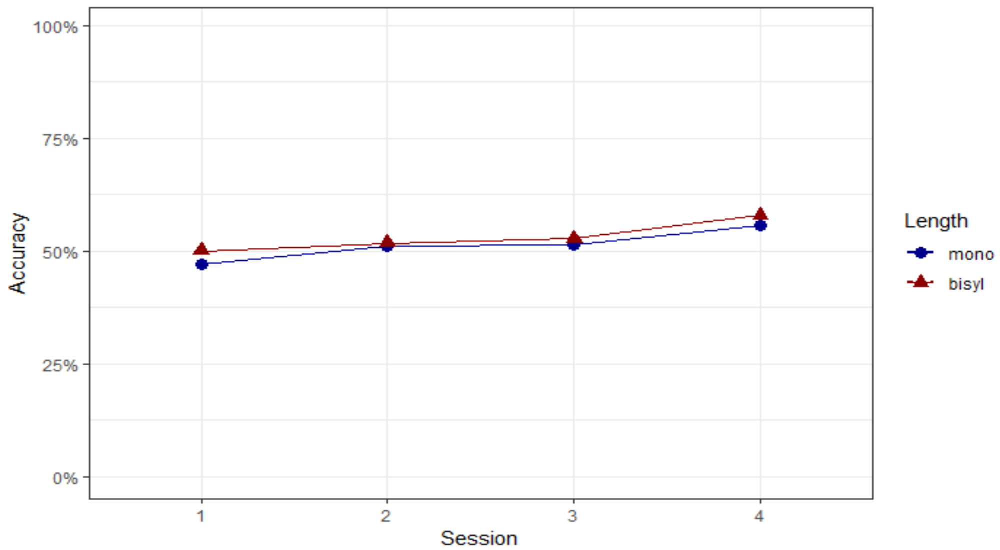

Figure 6-43. Forced-choice task: mean percentages of accuracy scores for interaction between session and word length condition: mono=monosyllabic words; bisyl=bisyllabic words.

Word Length	Session			
	1	2	3	4
Monosyllables	47%	50%	51.3%	55.4%
Bisyllables	50%	51.6%	52.5%	57.4%

Table 6-29. Forced-choice task: mean percentages of accuracy scores for interaction between session and word length condition. mono=monosyllabic words; bisyl=bisyllabic words.

6.2.4.3.1 Linguistic group, word length and session

A three-way ANOVA was conducted with Dprime as the dependent variable, word length and session as within-subject independent variables, and linguistic group as a between-subject independent variable to investigate whether or not an ability to detect new words depends on the interaction of word length condition, session, and linguistic group. Before the analysis, the data were aggregated according to Dprime score, subject, session, stress, and linguistic group variables. The results for the main effect in analysis of variance showed that this interaction was not significant [$F(3,72)=0.54$, $p=0.63$, $\eta_p^2=0.02$]. Pairwise comparisons for the linguistically sophisticated group showed that there were significant differences between the recognition of bisyllabic words in session 1 [$M=0.12$] and 4 [$M=0.56$], and between the recognition of monosyllabic words in session 1 [$M=0.36$] and 4 [$M=-0.02$], in both cases [$p<0.01$]. Despite a statistically insignificant main effect, this suggests that the recognition of both monosyllabic and bisyllabic words improved from session 1 to session 4 for the sophisticated group. None of the pairwise comparisons for the linguistically naïve group were significant; however, as Figure 6-44 illustrates there are trends indicating that performance on both bisyllabic and monosyllabic words improved from session 1 to session 4. The means of Dprime scores for the interaction between session, stress condition and linguistic groups are presented in Table 6-30.

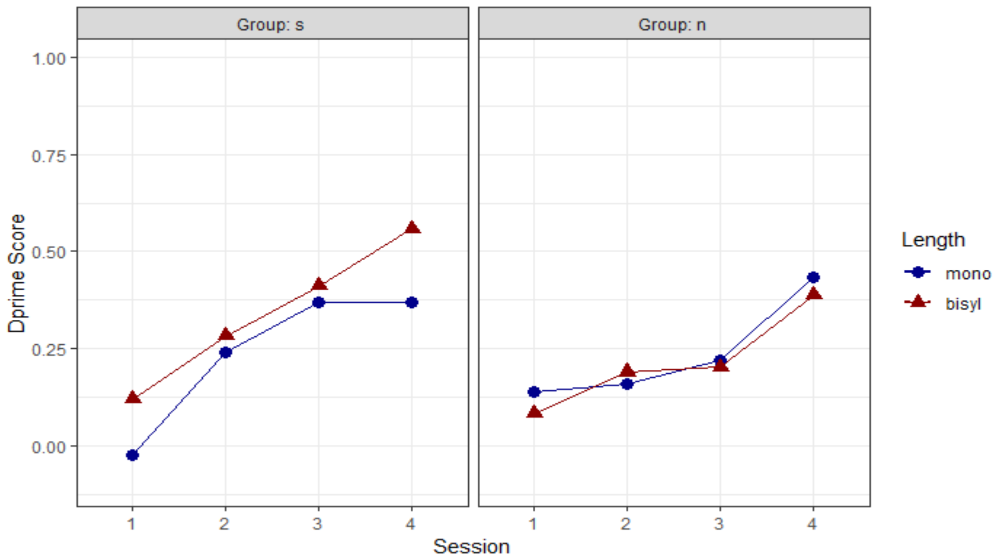

Figure 6-44. Word recognition task: Dprime scores for interaction between session, word length condition by linguistic group: s=sophisticated group; n=naïve group; mono=monosyllabic words; bisyl=bisyllabic words.

Session	Linguistically Sophisticated Group		Linguistically Naive Group	
	Length		Length	
	Monosyllables	Bisyllables	Monosyllables	Bisyllables
1	-0.02	0.12	0.14	0.08
2	0.24	0.28	0.16	0.19
3	0.37	0.41	0.22	0.20
4	0.37	0.56	0.43	0.39

Table 6-30. Word recognition task: Dprime scores for interaction between session and word length condition by linguistic group: s=sophisticated group; n=naïve group.

For the forced-choice task, a mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and session, word length, and linguistic group as explanatory variables. None of the model's results were statistically significant: (1) for the interaction between the naïve group, session 2, and bisyllables [OR = 1.14 (95% confidence interval [CI]: 0.59, 2.23), p=0.68]; (2) for the interaction between naïve group, session 3, and bisyllables [OR = 1.40 (95% confidence interval [CI]: 0.72, 2.70), p=0.32]; (3) and finally for the interaction between naïve group session 4 and bisyllables [OR = 1.40 (95% confidence interval [CI]: 0.72, 2.70), p=0.32]. The mean scores for this interaction are presented in Table 6-31, and Figure 6-45 below illustrates this interaction. The figure demonstrates that for both groups, there was an improvement in performance with further sessions for both levels of the word length condition. The performance on bisyllabic and monosyllabic words was similar for the linguistically sophisticated participants. Also, the performance of the sophisticated group for word length was better than that of the naïve group in session 4.

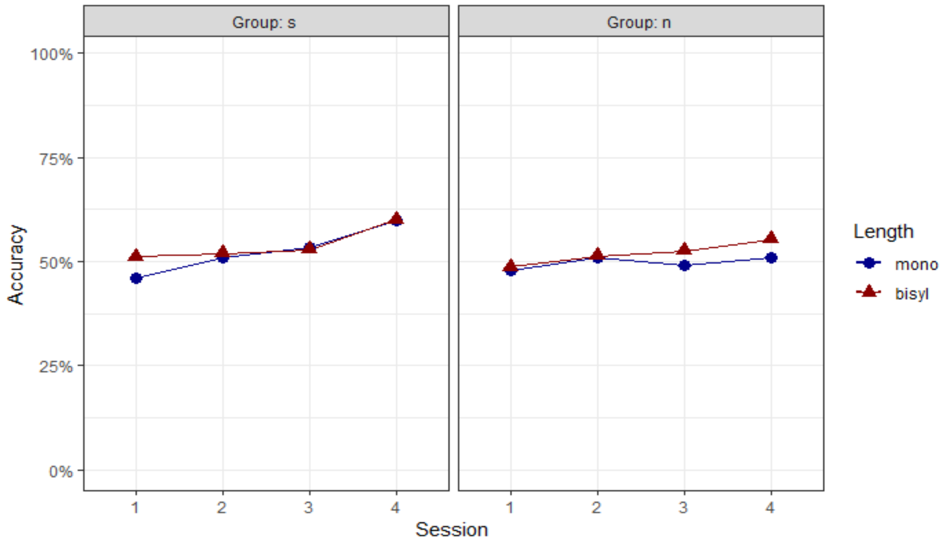

Figure 6-45. Forced-choice task: mean percentages of accuracy scores for interaction between session and word length condition by linguistic group: s=sophisticated group; n=naïve group.

Session	Linguistically Sophisticated Group		Linguistically Naive Group	
	Word Length		Word Length	
	Monosyllables	Bisyllables	Monosyllables	Bisyllables
1	46.2%	51%	48%	48.8%
2	50.8%	51.8%	51%	51.3%
3	53.3%	52.7%	49%	52.4%
4	59.4%	59.4%	51%	55%

Table 6-31. Forced-choice task: mean percentages of accuracy scores for interaction between session and word length condition by linguistic group: s=sophisticated group; n=naïve group.

6.2.4.3.1.1 Summary

The results for either the word recognition, nor forced-choice tasks yielded significant effects in terms of the better recognition of bisyllables than monosyllables, but there was a trend for accuracy to be higher in the recognition of bisyllabic compared to monosyllabic words. Also, the recognition of both bisyllabic and monosyllabic words improved over sessions. These results do not support *hypothesis* 9, which predicted that participants would be more accurate in detecting bisyllabic words than monosyllabic words, and that this ability would increase over sessions.

Moreover, there were no clear trends for the linguistic groups showing that performance for bisyllabic words was better than that for monosyllabic words

However, there was a tendency for the linguistically sophisticated group to have slightly higher accuracy on both monosyllabic and bisyllabic words than the linguistically naïve group. All in all, these results once again do not support *hypothesis 12*.

6.2.5 Generalisation

This section focuses on the analysis of the results in terms of whether or not participants could generalise according to the phonotactic properties they heard in the input. Recall that, in the word recognition task, to test participants' generalisation abilities, a type of stimulus variable was divided into targets and two groups of distractors (items with generalisable properties vs items with non-generalisable properties). To find out whether or not participants were generalising phonotactics of words which they heard in the input to novel words, it was hypothesised that accuracy (Dprime scores) for generalisable distractors would be lower than for non-generalisable distractors in the word recognition task. In order to calculate values of *d'* for the two groups of distractors, the type of stimulus variable was manipulated because the original design of this variable did not allow the calculation of *d'* as each level, which should have had targets and foils in order to calculate hits and false alarms. However, the level of target contained only targets (n=48) and so it was only possible to calculate hits and misses for this level of the variable, and the generalisable and non-generalisable levels contained only distractors (n=48 in each) and it was only possible to calculate false alarms and correct rejections. Therefore, a new variable of the type of stimulus was created with only two levels (of generalisable items vs non-generalisable items) with n=48 in each. Generalisable items, contained values of hits and misses from the level of targets in addition to its own rates of false alarms and correct rejections. Non-generalisable items also contained the same values of hits and misses from targets in addition to its own rates of false alarms and correct rejections. After this data manipulation, it was possible to calculate value of Dprime for generalisable and non-generalisable items. A one-way ANOVA was then conducted with Dprime scores as the dependent variable and the type of stimulus (generalisable vs non-generalisable items) as an independent variable to investigate whether or not *ab-initio* learners have generalisation abilities. Before the analysis, the data were aggregated according to Dprime score, subject and type of stimulus variables.

The analysis of variance revealed significant differences for the main effect of type of stimulus [$F(1, 27)=83.19$, $p< 0.001$, $\eta_p^2=0.75$]. The results show that the mean Dprime scores for non-generalisable distractors is 0.63, which is equal to about 70% accuracy, and the value for generalisable items is 0.19 which means that participants could discriminate these words only slightly better than by chance. Figure 6-46 below illustrates the mean Dprime scores across generalisable and non-generalisable items.

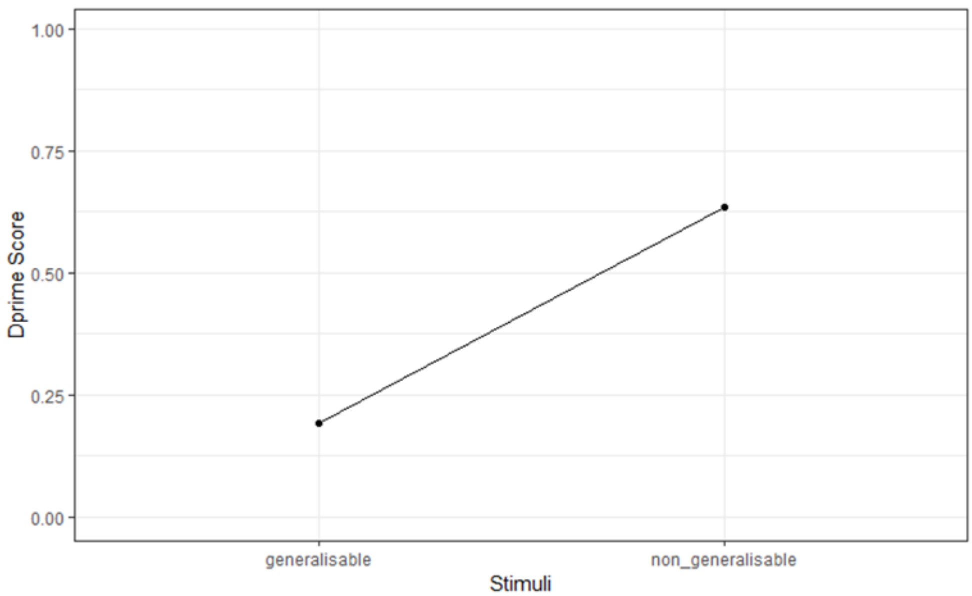

Figure 6-46. Word recognition task: Dprime scores across stimuli condition.

Furthermore, to investigate if participants were treating generalisable distractors like targets in the forced-choice task, the explanatory variable was type of stimulus (targets vs generalisable distractors), but it was no longer possible to use the dependent variable of accuracy which was used for the other analyses in the forced-choice task. The variable accuracy was no longer useful because, for targets, it represented percentage hits, and for generalisable distractors, it represented the percentage of correct rejections. Therefore, it made more sense to create a new dependent variable, which was *performance*. This was defined as the numbers of hits for targets and false alarms for generalisable distractors (that is, when participants thought that a distractor was a target). It was predicted that performance for targets and generalisable distractors would be similar because participants would incorrectly assume that a generalisable distractor was a target since both types of stimulus contained the same phonotactics. The mixed-effect logistic regression model contained this new variable of performance as the dependent variable and type of stimulus (generalisable items vs targets) as the explanatory variable. The results for the model indicate a statistically insignificant effect of type of stimulus. That is, for targets, the model produced the following results: [OR=1.25 (95% confidence interval [CI]: 0.91, 1.59), p=0.19]. The value of OR higher than one signals that the mean percentage accuracy was higher when the experimental stimulus was a target [M=55.7%] than when it was a generalisable distractor [M=51.4%]. Figure 6-47 illustrates the mean percentage performance for the type of stimulus.

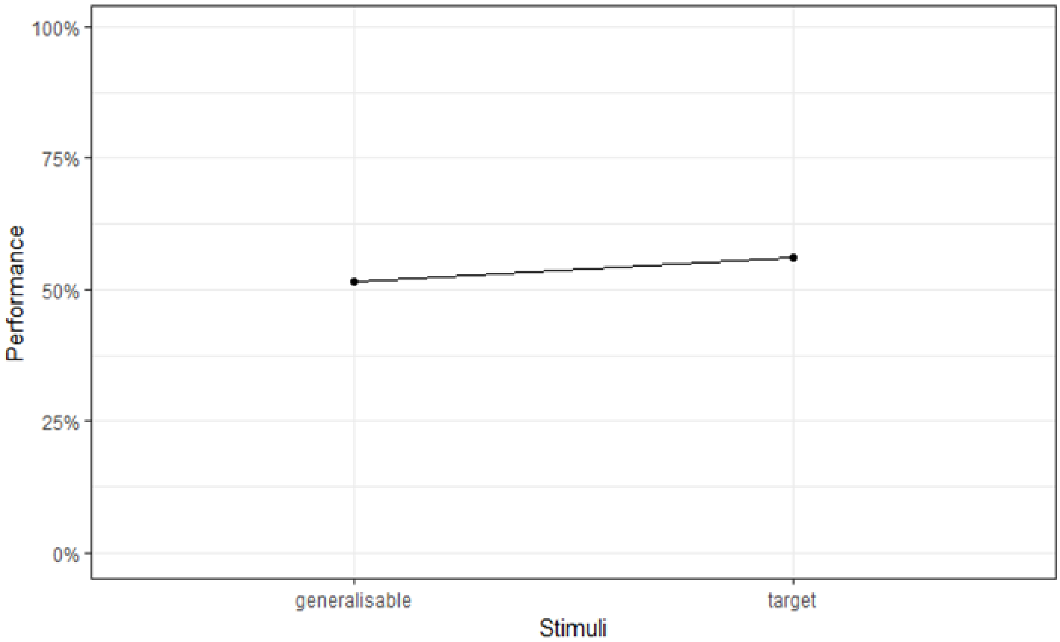

Figure 6-47. Forced-choice task: mean percentage performance for the type of stimulus.

As performance measure in the analysis above was defined differently for generalisable distractors and targets, one could argue that the results of the statistical test are not very credible. That is why, in addition to running a mixed-effect logistic regression, it was decided to calculate Dprime scores for types of stimuli in the forced-choice task. To calculate these Dprime scores, the numbers of hits with targets and false alarms with generalisable distractors were used. The results showed that the mean Dprime score is 0.11, with [min=-0.38] and [max=0.42]. These results confirm that discrimination between targets and distractors was truly only slightly above what would be expected by chance.

6.2.5.1 Linguistic group and generalisation

Finally, to test if there were differences between the linguistically sophisticated and naïve participants in their ability to generalise according to phonotactic properties heard in the input, the same analyses as in Section 6.2.5 were conducted for word recognition and the forced-choice task but where linguistic group was included as an additional between-subjects variable.

The results of the analysis of variance for the word recognition task showed that the interaction between type of stimulus and linguistic group was not significant: $F(1, 26)=0.22$, $p=0.64$, $\eta_p^2=0.002$. The pairwise comparison for the linguistically naïve group showed a statistically significant difference [$p<0.05$] between generalisable [$M=0.13$] and non-generalisable distractors [$M=0.60$]; and there was also a statistically significant difference between for the sophisticated group [$p<0.05$] between generalisable [$M=0.24$] and non-generalisable [$M=0.66$] distractors. These results indicate that there was no statistically significant difference between linguistically sophisticated and linguistically naïve groups since their recognition of generalisable distractors was merely a bit above chance, whereas the recognition of non-generalisable distractors corresponded to about 70% accuracy. Figure 6-48 below illustrates this interaction.

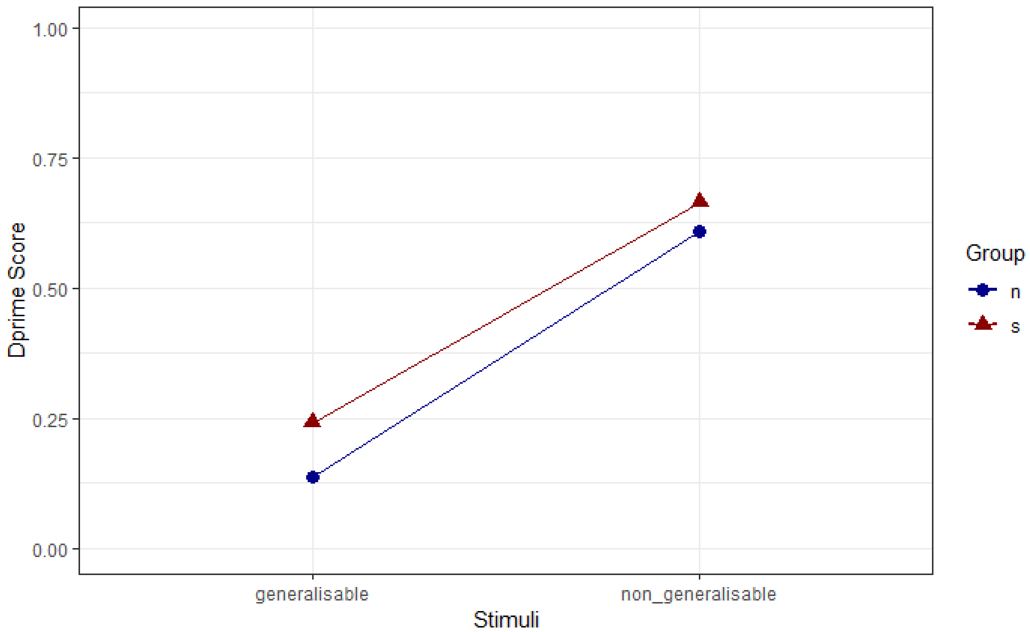

Figure 6-48. Word recognition task: Dprime scores for type of stimulus condition by linguistic group.

Results from the forced-choice task showed that the interaction between targets and the lingustically naïve group was not statistically significant with the following outcome: [$OR=0.86$ (95% confidence interval [CI]: 0.69, 1.07), $p=0.18$]. Despite these statistically insignificant results, performance on targets by the sophisticated group was higher [*M*=57.9%] than by the linguistically naïve group [*M*=53%]. Whereas performance on generalisable distractors was slightly above chance for both groups, at [*M*=51.9%] for the sophisticated group, and [*M*=50.7%] for generalisable distractors. Figure 6-49 below illustrates this interaction.

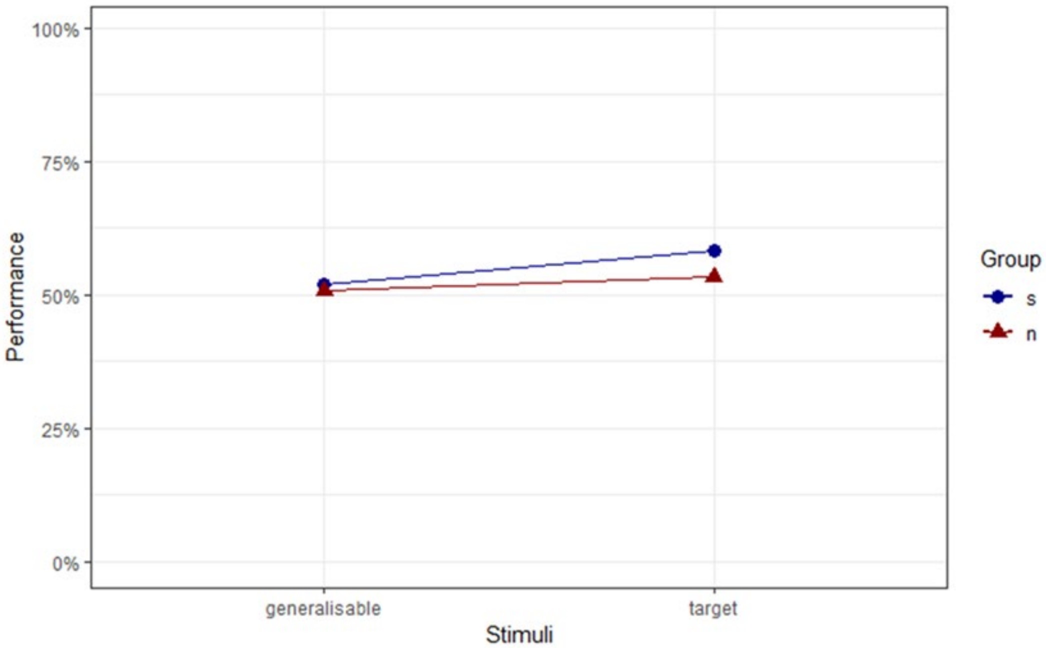

Figure 6-49. Forced-choice task: mean percentage performance for type of stimulus by linguistic group.

Additionally, Dprime scores were calculated for each linguistic group. As with the analysis of the main effect of generalisation, the number of hits from targets and the number of false alarms from generalisable distractors were used in order to calculate the Dprime scores. After that an ANOVA was run with Dprime scores as the dependent variable and linguistic group as an independent between-subjects variable. The results showed that the difference between the two groups in Dprime scores was not significant [$F(1, 26)=1.38$, $p=0.25$, $\eta_p^2=0.05$], although the mean Dprime score for the linguistically sophisticated group was slightly higher [*M*=0.15] than that for the linguistically naïve group [*M*=0.06].

6.2.5.1.1.1 Summary

To sum up, the results from the word recognition task showed that Dprime values for items with generalisable properties are significantly lower than for items with non-generalisable properties. This supports *hypothesis 10*, which predicted that Dprime scores for generalisable distractors would be lower than for non-generalisable distractors because participants would think that generalisable distractors were possible targets and non-generalisable distractors were not.

The results from the forced-choice task showed that participants incorrectly thought that generalisable distractors were targets, because the percentages of hits on targets and of false alarms on generalisable distractors were not significantly different, although the numbers of hits for targets were slightly higher than were false alarms for generalisable distractors. This was further confirmed by the Dprime scores, which indicated the existence of discriminability between targets and distractors slightly above chance levels. Consequently, the results from the forced-choice task can be said to support *hypothesis 11,* which predicted that performance on generalisable distractors and targets would be similar, given that both types of stimulus contained the same phonotactics.

Regarding performance according to linguistic group, the results for the word recognition task showed that there was no significant difference between naïve and sophisticated participants. This does not support *hypothesis 12,* which predicted differences between the two groups of participants. However, there was a trend for the Dprime scores of the linguistically sophisticated group to be higher than those of the linguistically naïve group for both types of distractors.

The results from the forced-choice task also show that there were no significant differences between the sophisticated and naïve groups. However, linguistically sophisticated participants received slightly higher hit rates for targets than linguistically naïve participants, whereas false alarm rates for generalisable distractors were a bit above chance levels for both groups. This was further confirmed by the fact that Dprime scores were not significantly different between the two groups, but the index of discriminability between targets and generalisable distractors was slightly higher among linguistically sophisticated participants. This does not support *hypothesis 13,* which predicted that that accuracy of the linguistically sophisticated group would be higher than that of the linguistically naïve group.

6.3 Cognate Identification Task

In the cognate identification task, participants were asked to listen to 20 words which were presented to them one by one. As with the word recognition task, participants needed to press the key 'z' on the computer keyboard if they thought they had heard the word in the input, and 'm' if they thought they had not heard it. Participants were asked to respond as quickly as possible, and if they failed to respond within four seconds then the program moved to the next item.

As the design of the cognate identification task was balanced, with 10 targets and 10 distractors, a mixed-effect logistic regression analysis of the data was used. The analysis of the results of cognate identification task started by determining the accuracy. Firstly, the numbers of hits, misses, false alarms and correct rejections were counted based on the design of the cognate identification task:

(1) Hit = if the type of stimulus was 'target', and response 'z';

(2) Miss = if the type of stimulus was 'target', and response 'm';

(3) False Alarm = if the type of stimulus was 'distractor', and response 'z';

(4) Correct Rejection = if the type of stimulus was 'distractor', and response 'm'.

Participants never failed to respond in this task, and so the response 'none' did not have to be taken into consideration. Hits and correct rejections were accurate responses; misses and false alarms were incorrect responses.

6.3.1 Paying attention to the input

To establish if participants had paid sufficient attention to the input, it first needed to be checked if there was an increase in accuracy scores over sessions. An increase in accuracy should be taken as an indication that participants had paid attention to the properties of the input. A mixed-effect logistic regression model was fitted with accuracy as the dependent variable, and session as independent variable, and subject and word variables were added to the model as random factors to investigate if accuracy in the recognition of cognates improved with more input. The results for the model indicated a significant effect of the session on day 2 [OR=1.37 (95% confidence interval [CI]: 1.04, 1.81), $p<0.05$], and a significant effect of the session on day 4 [OR=1.82 (95% confidence interval [CI]: 1.37, 2.43), $p<0.01$], the effect of session 3 was not significant [*OR*=1.1 (95% confidence interval [*CI*]: 0.84, 1.44), $p=0.50$]. The values of mean percentage accuracy for cognate words are presented in Figure 6-50 below. The results suggest that the mean percentage accuracy in session 1 was 62.8%, and this significantly increased from session 1 to session 2 [*M*=70%], and from session 1 to session 4 [*M*=75%]. However, the increase in accuracy scores from session 1 to session 3 [*M*=65.9%] was not statistically significant.

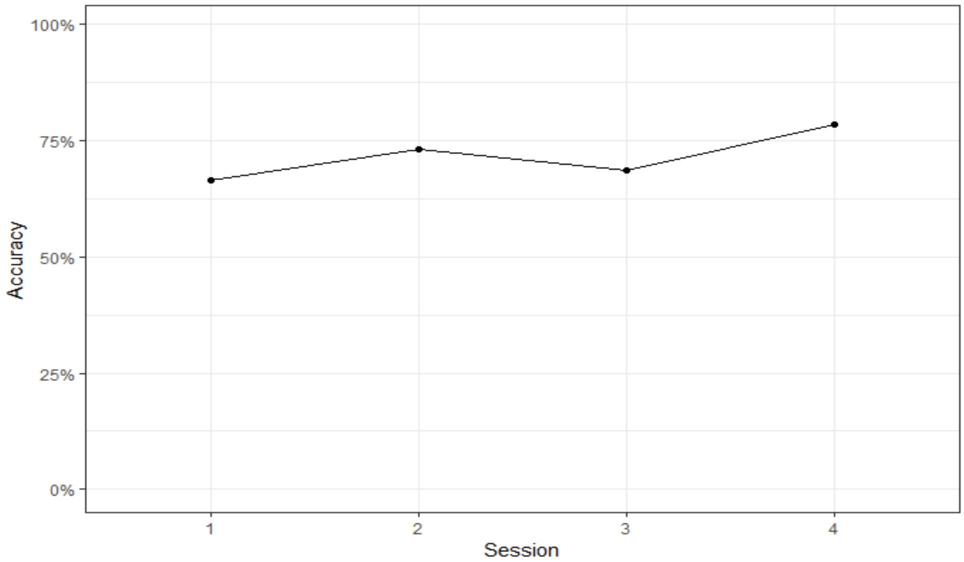

Figure 6-50. Cognate identification task: mean percentage accuracy scores across sessions.

In the design of this experimental condition, the audio recording which participants listened to in sessions 1 and 2 was different from the recording they heard in sessions 3 and 4. Unlike in the other two tasks within this experiment, this task utilised a set of new items which participants needed to respond to on sessions 3 and 4. Therefore, it made sense to ask if there was a statistically significant difference in accuracy between sessions 1 and 2, and between sessions 3 and 4. So, the session variable was converted into a new session variable, where sessions 1 and 2 were merged into a new level of session 1&2, and sessions 2 and 3 were merged into the second new level of session 3&4. The results for the model indicate a marginally significant effect of the new session variable [OR=1.19 (95% confidence interval [CI]: 0.98, 1.44), $p<0.08$] which shows that accuracy for cognate words in sessions 3 and 4 [M=70%] was slightly higher than accuracy in sessions 1 and 2 [M=67%].

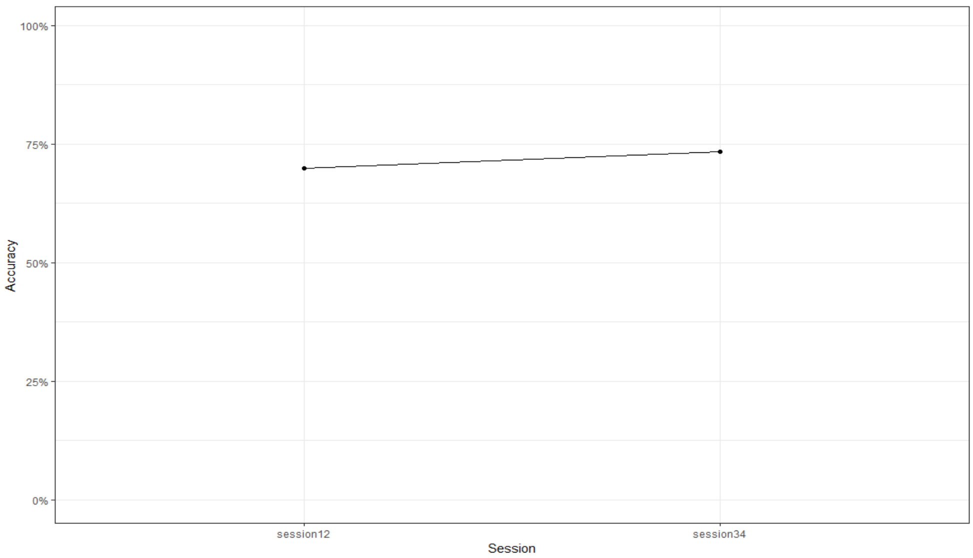

Figure 6-51. Cognate identification task: mean percentages of accuracy scores across new session variable.

6.3.1.1 Summary

To sum up, the overall results for the cognate performance task showed that there was a definite improvement in the detection of cognates over sessions. Recall that the aim of this task was to make sure that participants paid attention to audio-recordings throughout all sessions, and to eliminate any outlier participants. To identify any such outliers, mean percentage accuracy scores for each participant were calculated over the four days to give an overall score. These scores are presented in Table 6-29 below. The minimum accuracy score is [*min*=41.2%], and the maximum is [*max*=90%]. The median of these scores is [*median*=68.1%], and the mean is [*mean*=68.95%]. In statistics, the interquartile range (*IQR*) is commonly used to establish the spread of observations in a dataset and, technically speaking, an outlier is any value which is distant by 1.5 times above a higher *IQR* or below a lower *IQR* in this dataset. The third quartile in Table 6-32 is [*Q3*=76.53%], and the first quartile is [*Q1*=62.17%], and the *IQR*=14.35%. The lower range limit was calculated to be 40.67% and the higher range limit which 98%. It is clear from inspecting Table 6-32 that none of the participants fell below the lower limit exceeded the higher limit, and so none of the subjects were excluded from the analysis due to their inability to pay attention.

Number	Participant ID	Accuracy (mean %)
1	Part26	41.2%
2	Part16	51.2%
3	Part07	55%
4	Part28	56.2%
5	Part21	60%
6	Part18	61.2%
7	Part10	61.2%
8	Part23	62.5%
9	Part19	66.2%
10	Part20	66.2%
11	Part12	66.2%
12	Part11	66.2%
13	Part15	67.5%
14	Part13	67.5%
15	Part05	68.7%
16	Part03	70%
17	Part04	71.2%
18	Part02	73.7%
19	Part14	73.8%
20	Part17	76.2%
21	Part06	76.2%
22	Part24	77.5%
23	Part01	78.3%
24	Part09	78.8%
25	Part25	81.7%
26	Part08	82.5%
27	Part22	83.6%
28	Part27	90%

Table 6-32. Cognate identification task: mean percentages of accuracy score for each participant.

6.4 Overall summary

The results from the word recognition, forced-choice and cognate identification task have been presented in this chapter. Additionally, a summary of signal detection theory (Green & Sweets

1966) was provided along with a rationale for its choice for the analysis of the data from the word recognition task. The motivation behind all statistical tests selected for the testing of hypotheses was also explained.

A summary of the main findings for each hypothesis is given below. Firstly, the results are discussed for those hypotheses which were tested and we supported using both the word recognition and forced-choice tasks. Secondly, the results are considered for those hypotheses which were tested using both tasks but were supported by only one task. Thirdly, the hypotheses which were tested using both tasks, and were not supported by either but where the results were deemed to be significant. Fourthly, hypotheses which were tested by both tasks, and were supported by neither task because the results were not significant are considered. Then, the results are summarised for hypotheses which tested generalisation and the results comparing the linguistic groups. Finally, I discuss the results for the cognate identification task.

1. Hypotheses which were tested using both the word recognition and forced-choice tasks, and were supported by the results for both tasks

Hypothesis 1 predicted that participants' accuracy on targets would improve with an increased amount of input. This was supported by the results of the word recognition task where accuracy (Dprime scores) significantly increased in each session except from session 2 to session 3. However, it was only partially supported by the results of the forced-choice task because accuracy (percent correct) on targets on session 3 was 52.1%, which was marginally higher than 49% on day 1, but there was a significant increase in accuracy scores on session 4 at 56.7%.

2. Hypotheses which were tested using both the word recognition and forced-choice tasks, but were supported by only one task

Hypothesis 2 predicted that participants in this study would be more accurate in detecting targets which have native phonotactics than non-native phonotactics. Results from the word recognition task showed that accuracy (Dprime scores) on targets with native phonotactics was significantly higher than on targets with non-native phonotactics. However, the results from the forced-choice task did not find significant differences in accuracy (percent correct) between target words with native and non-native phonotactics as words were recognised only slightly above chance levels regardless of phonotactic pattern. Interestingly, there was a trend in the results of the forced-choice task showing performance for words with non-native phonotactics was slightly better than for words with native phonotactics.

3. Hypotheses which were tested using both the word recognition and forced-choice tasks and were not supported, but the results were deemed to be significant

Hypothesis 3 predicted that participants in this study would be more accurate in detecting targets which were stressed on the first syllable. This was not supported by the results from either task. However, the results of the forced-choice task showed that accuracy (percent correct) on words stressed on the second syllable was marginally higher, at 57.2%, than on words with a strong initial syllable at 48.6%. The results of the word recognition task were not statistically significant, but there was the same trend that accuracy (Dprime scores) on weak-strong words at 0.30, was slightly higher than that on strong-weak words at 0.25.

Hypothesis 5 predicted that participants in this study would be more accurate in detecting targets with native phonotactics and which were stressed on the first syllable, than detecting targets with native phonotactics and word-final stress. Results from the word recognition task and the forced-choice task did not support this hypothesis. This is because the results from both tasks showed that words which followed native phonotactics and were stressed on the final syllable received the highest accuracy (Dprime for the word recognition task and percent correct for forced-choice task), and they were recognised significantly better than words with native phonotactics and word-initial stress, which is the opposite of what the hypothesis predicted. Additionally, when words followed non-native phonotactics, there was higher accuracy on strong-weak words rather than weak-strong words and this pattern was common across the two tasks.

4. Hypotheses which were tested using both the word recognition and forced-choice tasks, and were supported by neither task because the results were not statistically significant

Hypothesis 4 predicted that participants of this study would be more accurate in detecting bisyllabic targets than monosyllabic targets. This was supported by the results of neither the word recognition task nor the forced-choice task. However, there was a minor trend that accuracy on bisyllabic words was slightly higher than on monosyllabic words.

Hypothesis 6 predicted that participants of this study would be more accurate in detecting targets which have native phonotactics rather than words with non-native phonotactics, and there should have been a preference for bisyllabic than monosyllabic targets. This hypothesis was not supported by the results for either task. However, there was a slight preference for bisyllabic words over monosyllabic words when words followed native phonotactics, and there was a minor preference for monosyllabic words over bisyllabic words when phonotactics were non-native. This trend was common to both tasks.

Hypothesis 7 predicted that participants in this study would be more accurate in detecting targets which have native phonotactics than targets with non-native phonotactics, and this ability would increase over sessions. This hypothesis was not supported by the results of either task. However, there was a trend for the accuracy for words with native phonotactics to be slightly higher than for non-native phonotactics in the word recognition task. Whereas, the trend was the opposite, so that accuracy was somewhat higher on words with non-native phonotactics than with native phonotactics, in the forced-choice task.

Hypothesis 8 predicted that participants in this study would be more accurate in detecting targets which are stressed on the first syllable than on the second syllable, and this ability would increase over sessions. This hypothesis was supported by the results of neither task. However, there was a trend common to both tasks in that accuracy was higher on weak-strong words than on strong-weak words

Hypothesis 9 predicted that participants in this study would be more accurate in detecting bisyllabic targets than monosyllabic targets, and this ability would increase over sessions. This was also supported by neither task. However, there was a trend that bisyllabic words were recognised slightly better than monosyllabic words.

5. Hypotheses about generalisation

Hypothesis 10 predicted that Dprime scores involving generalisable distractors would be lower than those for non-generalisable distractors because participants would think that generalisable distractors were possible targets, and that non-generalisable distractors were not. This was tested only by the results of the word recognition task, and it was supported.

Hypothesis 11 predicted that performance (measure in percent correct for number of hits for targets and number of false alarms for generalisable distractors) would be similar between targets and generalisable distractors provided that both types of stimulus contained the same phonotactics, because participants would incorrectly think that generalisable distractors were targets. This hypothesis was tested only by the results of the forced-choice task, and it was supported.

6. Hypothesis about linguistic groups

Hypothesis 12 predicted that linguistically sophisticated participants were expected to have higher accuracy or performance than linguistically naive participants with respect to each of hypotheses 1-11 discussed above. This was not supported for all hypotheses. That is, the difference between linguistically sophisticated and linguistically naïve participants was not statistically significant in most cases. However, with respect to all of the findings discussed above, the accuracy or performance of linguistically sophisticated participants was higher than that of linguistically naïve participants. Additionally, with respect to *hypothesis 2*, which predicted that participants in this study would be more accurate in detecting targets which have native phonotactics than non-native phonotactics, the forced-choice task showed that linguistically sophisticated participants had nearly the same accuracy on words with native phonotactics at 53.5% and non-native phonotactics at 53.1%, whereas linguistically naïve participants had higher accuracy on words with non-native phonotactics at 54% and words with native phonotactics were recognised below chance levels at 47%.

Finally, the results from the cognate identification task showed that there was a significant increase in accuracy in all sessions when they were compared with the first session except for session 3. However, the values of mean percentage accuracy were slightly higher for sessions 3 and 4 when participants were tested on cognates about 'university life' than compared to sessions 1 and 2 when they were tested on cognates about 'music'. Additionally, no participant was excluded on the basis of paying too little attention to the input.

A discussion of these results is presented in the next chapter.

Chapter 7. Discussion

7.1 Introduction

The purpose of this chapter is to discuss the results of the present study. Section 7.2 discusses the results obtained from the word recognition and forced-choice tasks. It comprises several sub-sections reflecting on each research question listed in Section 5.4 except for the last one about the difference between linguistically sophisticated and linguistically naïve participants – it refers to all the previous questions and is discussed separately in Section 7.3. Section 7.4 discusses the results of the cognate identification task. Finally, a summary of this chapter will be presented in Section 7.5.

7.2 Discussion of results on the word recognition and forced-choice tasks

The results on the word recognition and forced-choice tasks will be discussed together because the purpose of both tasks was to see if participants could detect words that they heard in the input as opposed to words they did not hear, and to examine which cues participants relied on for word detection. Additionally, both tasks had another purpose, i.e. to investigate if learners could generalise to phonotactic properties heard in the input.

While discussing the results, it is important to remember that the word recognition and forced-choice tasks had different designs: In the word recognition task, after listening to the input, participants were asked to listen to 144 words containing targets as well as generalisable and non-generalisable distractors, and for each word, they had to decide whether they had heard this word in the input. In the forced-choice task, after listening to the input, participants were asked to listen to 48 pairs of words, where each pair comprised a target and a generalisable distractor.

Participants had to decide which of two words in each pair they had encountered in the input. Both tasks were timed, with the response to be given within four seconds.

7.2.1 Effect of session

The 1st research question asked if an ability to detect Russian words from the input would increase over sessions. The word recognition task and forced-choice tasks found the effect of session but to different extents. In particular, the results of the word recognition task showed that there was a significant difference in participants' ability to detect words from the input between all sessions except the second to the third session. In contrast, the results from the forced-choice task showed only a marginally significant effect of the third session and a significant effect of the fourth session when compared with the first session. However, what was common between both tasks is that there was a clear improvement trend in accuracy scores throughout all sessions. It is consistent with other studies on *ab-initio* learners which showed that increasing the amount of input positively correlated with improved accuracy (Rast & Dommergues 2003; Rast 2008, 2010; Gullberg *et al.* 2010, 2012; Shoemaker & Rast 2013; Carroll 2014).

Gullberg *et al.* 2012 found an effect of segmentation of frequent words (eight occurrences in the input) after as little as seven minutes of exposure. The present study did not specifically aim to investigate the effect of frequency, but it is very much consistent with the results of Gullberg *et al.* 2012, albeit a different experimental setup was used. Each target occurred exactly once in about three and a half -minute audio recording that participants listened twice before testing, leading to seven minutes of exposure to Russian per session. Hence participants heard all target words in the input exactly twice. In the present study, a significant effect of session was observed already in the second session for the word recognition task, where the Dprime was used as the measure of accuracy. By the end of the second session, participants of the present study accumulated about 14 minutes of exposure to Russian, and the target words appeared four times in the input (vs seven minutes and two occurrences of target words after the first session). Despite the lower frequency of target words in the present study that participants received by the end of the second session, i.e. four occurrences per 14 minutes vs eight occurrences of "frequent" words per seven minutes in Gullberg *et al.* (2012) and lower ratio of "frequent" to "infrequent" words, i.e. 4/2=2 in the present study after second / first sessions, or 8/2=4 in Gullberg *et al.* (2012), frequent words were identified significantly better than infrequent words in both the present study and Gullberg *et al.* (2012). Note that there was another interesting difference in the setup of the experiment: In the present study, the same words were used across several sessions and the frequency or, more precisely, the total amount of exposure was increasing during each session, whereas Gullberg *et al.* (2012) used different words within a single session.

As a matter of fact, Gullberg *et al.* (2012) found that accuracy of a single exposure group (after seven minutes of exposure or eight times exposure to frequent words) was 55% which closely corresponds to the Dprime value on the second day (*d'*=0.36). It means that in the present study, accuracy on targets was above chance even on the first day (*d'*=0.20, i.e. positive). Moreover, the present study shows that the ability to detect words also significantly increased between the third and fourth sessions: Participants were about 65% (*d*'=0.57) accurate on spotting target words among distractors on the final day. Once again, it shows that the results of the present study are consistent with those of Gullberg *et al.* (2012) which found that the accuracy of double exposure group (when frequent words occurred 16 times) was 60%. The participants of the present study heard exactly eight instances of targets in the input by the end of the fourth session. However, it needs to be explained why the accuracy in the word recognition task in the present study is slightly higher than that in Gullberg *et al.* (2012). It is feasible that it is because the present study utilised *within-subject design* (meaning that all participants took part in all sessions), whereas

Gullberg *et al.* (2012) utilised *between-subject design* (meaning that half of the participants were in the single exposure group and another half were in the double exposure group). It is commonly acknowledged that both types of experimental designs have their merits and the choice of one over the other should be carefully decided based on research questions and practicalities of a study, see for discussion Charness *et al.* (2013: 2). However, there is one disadvantage of the within-subject design, which needs to be discussed, which is the carryover effect, i.e. when all participants get tested just once, it can affect all the subsequent testing in undesirable ways due to accumulating practice. Instead of counterbalancing, which is a common measure to avoid effects of practice and fatigue in within-subject designs (Allen 2017), the present study utilised different inputs, that is on the first and second days participants listened to input sentences which contained cognates about music, whereas they listened to different input sentences which contained cognates about university life on the third and fourth days. However, it is conceivable that the practice which participants received during the testing (i.e. participants were tested on the same word recognition task and the forced-choice task on each of four sessions) presented participants with an additional input to the one participants received during the listening to the input phase on each day, which could have resulted in the higher accuracy that was observed in the performance of the word recognition task in the present study than those of Gullberg *et al.* (2012).

Interestingly, one of the findings of the present study concerning the first research question is that the effect of the session was much more pronounced in the word recognition task than in the forced-choice task: For the former, the effect is present already during the second session, as well as between the third and fourth sessions. In contrast, the first significant effect of targets' identification was found only on the third session in the forced-choice task, with this effect being only marginally significant. These could be due to the forced-choice being more *difficult* than the word recognition task because it involved the choice between two words where one word was a target but another one was a distractor which matched with a target in phonotactics, word length, and stress for bisyllabic words, e.g. the target [klɨk] was paired with the distractor [klat], the target [ˈklʲevʲɪr] was paired with the distractor [ˈklapən], and the target [klaˈtʃʲok] was paired with the distractor [klʊˈbok]. Hence, the word recognition task was more manageable than the forced-choice task because in the word recognition task participants responded according to "feel" on every item which was presented one by one. However, the reason is uncertain, e.g. one may instead argue that recognition of targets could be easier in the forced-choice task as participants know that one word in a pair has to be a target.

Nevertheless, given that the first marginally significant effect of accuracy was observed on the third session (M=52.1%) which was just slightly above chance level, and a significant effect of the fourth session (M=56.7%) while the accuracy on the first session was just slightly below chance at 49%, we can indeed conclude that identification of target words took place despite the difficulty of deciding between two very similar stimuli.

The findings from the word recognition and the forced-choice task provide evidence that learners' ability to detect words from the input does indeed increase over sessions, with this ability starting to appear only on the third session in the forced-choice task. These dissimilar results are interesting but perhaps are not surprising if, in addition to what was already discussed about each task, we consider what underlying abilities are tapped by these tasks. In particular, the word recognition task is an implicit memory task because it tested participants' responses to language stimuli without their awareness and automatically as each experimental stimulus was presented to participants individually and they needed to respond within four seconds whether they heard it in the input before. In contrast, the forced-choice task was more of an explicit memory task because it required participants to make a conscious decision within four seconds about which of two very similar sounding words appeared in the input phase[50]. Consequently, performance on these tasks presupposed the involvement of *implicit knowledge* in the word recognition task and *explicit knowledge* in the forced-choice task. According to R. Ellis (2009: 3-6), implicit knowledge is usually gained through implicit learning without demands on working memory, and it results in knowledge which cannot be verbalised; whereas explicit knowledge can either be a product of implicit or explicit learning but there is evidence that learners are aware of this knowledge. Moreover, implicit knowledge usually proceeds to explicit knowledge. There are researchers who disagree that implicit and explicit processes should be dissociated (e.g. Doughty 1991; Shook 1994), but most would agree that these are different processes (e.g. N. Ellis 1993; Robinson 1996; Norris & Ortega 2000; DeKeyser 2003; Hulstijn 2005; Gass & Selinker 2008; R. Ellis 2009). It will be shown next how the different performance in the word recognition task and the forced-choice task could potentially be explained in the light of implicit and explicit knowledge. R. Ellis (2009: 13) suggests that "difficulty in performing a language task may result in the learner attempting to exploit explicit knowledge". As was already mentioned above, the forced-choice task in the present study was more difficult as it required participants to make a conscious decision wither it was the first or the second word which they heard during the listening phase. The accuracy in the forced-choice task on the first day was 49%, i.e. slightly below chance level, and it was slowly increasing, becoming marginally significant at the third session and significant at the fourth session. That is, at the first session the participants found the task too demanding, but gradual improvement in the accuracy could indicate that the participants started to draw on explicit knowledge to accomplish the task. In contrast, the accuracy in the word recognition task

[50] Some may argue that the forced-choice task in the present study actually was not an explicit knowledge task because the task was time-pressured, while explicit knowledge tasks are normally do not have time constraints. Therefore it was said that the forced-choice task is more of an explicit memory because it ticked other criteria which an explicit memory task should have (see R. Ellis 2009: 40), such as the task encouraged participants to respond using 'rules' but not in accordance to 'feel' and to respond using metalinguistic knowledge.

was already above the chance level at the very first session, which might suggest that participants were already exploiting implicit knowledge. It means that participants were already exhibiting the earliest sensitivity to the isolated word forms which were extracted from the sequential context they heard during the familiarisation phase, in the absence of conscious learning effort. Moreover, since implicit knowledge precedes explicit knowledge, the significant improvement in the accuracy in the word recognition task was observed earlier than in the forced-choice task. The results on other research questions of the present study are discussed in the light of implicit and explicit knowledge in the following sub-sections.

Before we move to the discussion of results on the effect of individual cues for detection of Russian words, recall that both word recognition and forced-choice tasks showed no significant differences in detection of words between the second and third sessions, because participants were exposed to the new input which contained the same targets embedded in new sentences containing cognates about the university life rather than music. These results can be taken as evidence that change of input influenced participants' ability to detect words as accurately as they could if input was not changed because, at the fourth session after listening to the same input as in the third session, participants were showing significant improvement again.

7.2.2 Effect of single cues (phonotactics, stress, word-length)

7.2.2.1 Effect of phonotactic cues

The 2nd research question of the present study asked whether learners rely on L1 phonotactics, or they develop sensitivity to Russian phonotactics when detecting words of Russian from the input. In fact, it was predicted that *ab-initio* learners would rely more heavily on their knowledge of L1 (English) phonotactics. The present study found that *ab-initio* learners were indeed relying more on their knowledge of L1 phonotactic constraints than Russian phonotactics when detecting words of Russian in the word recognition task. This is consistent with the previous research on the effect of L1 phonotactic cues for speech segmentation. In particular, psycholinguistic tasks studies which measured the on-line performance of proficient bilinguals who appeared to be activating their knowledge of L1 phonotactics when listening to an L2 indicating element of L1 transfer (Altenberg & Cairns 1983; Weber 2000; Weber & Cutler 2006). Moreover, the findings of the present study in the word recognition task support the existing studied on *ab-initio* learners (Rast & Dommergue 2003; Rast 2008) which showed that learners were more accurate in repeating L2 words if they contained a segment or a cluster that existed in their L1, suggesting effect of L1 transfer on L2.

Contrary to the word recognition task, the results of the forced-choice task showed that there was no significant difference between words with L1 English and L2 Russian phonotactics, but words with Russian phonotactics were slightly preferred over words with English phonotactics with latter being recognised nearly at a chance level. This result, unlike the result of the word recognition task, does not provide evidence for the effect of the transfer of L1 English phonotactics, but also it does not show that participants could rely on L2 Russian phonotactics, although there are indications of emerging sensitivity.

The fact that the result of the word recognition task support the prediction that learners rely on L1 phonotactics, but the result of the forced-choice task does not, can be analysed in the light of implicit and explicit knowledge. As discussed in the previous sub-section, participants could be drawing on explicit knowledge in the forced-choice task and implicit knowledge in the word recognition task. As there was no difference between L1 English and L2 Russian phonotactics, but there was a slight preference for Russian phonotactics in the forced-choice task, participants might have invoked a conscious strategy of accepting the words which sounded “least English”, thus selecting more targets with non-English phonotactics. This can be explained by *psychotypology* (Kellerman 1979), which proposes that language learners have their perceptions about differences and similarities between source and target languages, and these can affect choices made in those languages. Kellerman (1979 cited in Gass and Selinker 2008: 138) notes that L2 learners may be sceptical about similar structures in the source and target languages, which can make them avoid using these structures and focus their attention on what is different between two languages. For example, in the present study native English speakers could have noticed that clusters with MSD=0 (e.g. *kn*-, *zv*-, *tv*-) do not "sound right" in English, so they must belong to the target language. Moreover, participants could have noticed that the target language which they were exposed to sounded like as if it belonged to the Slavic languages group. The study by Skirgard *et al.* (2017) utilised a large sample of participants from all over the world who needed to listen to a clip of a speech from 78 different languages in the online Great Language Game and to guess which language it was in the multiple-choice. Among other results, the study showed that there was much confusion in deciding among Slavic languages because participants mistook one language for another. Another study demonstrated that naïve listeners showed above chance sensitivity to differences between German and Russian after listening to these two languages recorded by the same speaker (Kirk *et al.* 2013). The tasks in both studies were not timed, and participants could take as much time as they needed to listen to language extracts, which means that participants there too, they were relying on their explicit knowledge.

As already mentioned above, contrary to the forced-choice task, the word recognition task was less conscious, i.e. participants did not have a chance to become aware of their choices; therefore, the effect of the L1 transfer of English phonotactics was present.

Effect of stress cues

The 3rd research question of the present study asked whether *ab-initio* learners rely on the strong-weak stress pattern (also known as MSS), see Cutler 1990, Cutler 1994) or on the weak-strong stress pattern when detecting words from the input. There was a common trend between word recognition and forced-choice tasks that showed that words stressed on the second syllable (weak-strong pattern) were recognised better than words which were stressed on the first syllable (strong-weak pattern). However, the result from the forced-choice task was marginally significant, whereas the result from the word recognition task was not significant. These results mean that the present study did not find the effect to support MSS. The opposite (weak) effect was observed in the present study, that is participants were more accurate in detecting words which were stressed on the second syllable, e.g. [klaˈtʃʲ ok] was preferred over [ˈklʲevʲɪr]; or [knʲaˈzʲok] was preferred over [ˈknʲigəm]. This finding is somewhat surprising, as it is not consistent with a number of previous studies which found a strong role of MSS when detecting words in L1 speech segmentation (e.g. Cutler & Norris 1988; Cutler 1990; Cutler & Butterfield 1992; Jusczyk *et al.* 1993b; Cutler 1994; Turk *et al.* 1995) and in L2 English speech segmentation (e.g. Archibald 1992, 1993). This could be due to significant differences between the stress placement in Russian and English.

As discussed in Chapter 4, researchers generally agree that Russian stress placement cannot be predicted by an underlying rule and the stress can be placed on any vowel within a word and is a part of knowledge about a word. As Russian allows bisyllabic words being stressed either on the first syllable or the second, the present study did not predict *ab-initio* learners' sensitivity to Russian stress. It was only predicted that *ab-initio* learners would follow MSS when detecting words in Russian, as indeed there is substantial evidence that English learners do so when segmenting their native language English.

Although the stress placement in Russian has been greatly debated among linguists, some researchers take strong positions by claiming that Russian has an iambic foot (Halle & Vergnaud 1987; Melvold 1990; Alderete 1995; Crosswhite 2001) meaning that stressing a word-final syllable in multiple syllables words is more common than stressing the first syllable. For example, the study by Crosswhite *et al.* (2003) asked native speakers of Russian to read sentences in Russian where one word in each sentence was a nonsense trisyllabic word generated according to Russian phonotactics and taking a noun position within a sentence. The main results of the study showed that bare or non-morphemic words, e.g. [nav^{j}ekum], were stressed on the final syllable at 90%, followed by stress on *medial* syllable at 9%, and only 1% of words were stressed on the first syllable. In contrast, suffixed or morphemic words, e.g. [b^{j}atʃ^{j}el^{j}-am] where [b^{j}atʃ^{j}el] is a stem and [-am] is a highly productive dative plural morpheme, were stressed more than 70% on the medial stress, about 20% on the final stress, and about 10 % on the first syllable. Researchers took these results as evidence that stress in Russian is placed on the final vowel of a stem, suggesting an iambic foot. Though the study of Crosswhite *et al.* (2003) were criticised by Mołczanow *et al.* (2013), the results of Crosswhite *et al.* (2003) seem very plausible. Hence, the preference of weak-strong words over the strong-weak words in the forced-choice task of the present study is not so surprising and may indicate that participants were showing sensitivity to the weak-strong stress when detecting Russian words.

Another explanation could be obtained if the results of the present study are once again analysed in the light of implicit vs explicit knowledge. Indeed, the sensitivity to the weak-strong stress pattern was marginally significant only in the forced-choice task, which draws on explicit knowledge. Participants could have adopted a strategy of selecting words which sound “least English” and prefer words with word-final stress, which also fits with psychotypology (Kellerman

1979) as discussed in the previous sub-section. Note that this explanation is orthogonal to the one in the previous paragraph, and both could have contributed to the observed effect.

7.2.2.2 Effect of word-length cues

The 4th research question of the present study asked whether *ab-initio* learners show a preference for bisyllabic over monosyllabic words when detecting words from the input. The results of both the word recognition and forced-choice tasks were not significant, i.e. there is no conclusive evidence that participants used the word length (measured in the number of syllables) as a cue. In particular, though bisyllabic words were recognised better than monosyllabic words, the result was not significant. This finding is surprising because word-length cue was shown to be important in statistical studies on artificial speech segmentation (e.g. Saffran *et al.* 1996a,b; Aslin *et al.* 1998), as well as segmentation of L1 English language by infants in the study by Johnson and Jusczyk (2001). Furthermore, word length cue was important in studies on word recognition by *ab-initio* learners. For instance, Gullberg *et al.* (2012) found that Chinese bisyllabic words were recognised significantly better than monosyllabic words by Dutch native speakers. Also, there were higher success rates with longer than shorter words in a study by Carroll (2014), who found a better recognition of words comprising of six syllables than on words of four and five syllables. Moreover, Rast (2010) found that longer words with three to six syllables were better translated that shorter words.

The finding of the present study that Russian bisyllabic words were not recognised significantly better than monosyllabic words appear to add to the findings by Rast and Dommergues (2003) and Rast (2008) who found no effect of word length (measured from 1 to 6 syllables) on French L1 learners' ability to repeat Polish words even after eight hours of input. However, the comparison of the result of the present study with that of Rast and Dommergues (2003) and Rast (2008) should be taken with caution as while the present study utilised psycholinguistic tasks which investigated perception of English native speakers of Russian, those studies asked participants to repeat target words after hearing them in sentences, therefore the effect which they observed is likely due to production constraints. As the present study found trends of bisyllabic words being recognised slightly better than monosyllabic words, it points to some evidence of learners' sensitivity to longer words. A study by Dommergues and Segui (1989) found that monosyllabic, rather than bisyllabic words presented problems in processing. It could be that the bisyllabic words of the present study were not long enough for the difference to be significant. It is an interesting research question, and one would need to put it to test to see if, for instance, three-syllable Russian words would be recognised better than bisyllables and/or monosyllables. The results on the interaction between phonotactics and stress, and between phonotactics and word length are discussed in the next section.

7.2.3 Effect of combination of cues

7.2.3.1 Effect of combination of phonotactics and stress

The 5th research question of the present study asked whether *ab-initio* learners guided by an interaction between phonotactics and MSS when detecting novel words. Moreover, it was predicted that *ab-initio* stage learners would be more accurate in detecting targets with native phonotactics and which are stressed on the first syllable due to the robust effect of L1 transfer on L2 segmentation (Altenberg & Cairns 1983; Weber 2000; Weber & Cutler 2006) and MSS (Cutler & Norris 1988; Cutler 1990; Cutler 1994).

The results of the present study were significant in both word recognition and forced-choice tasks which showed that participants were detecting best of all target words which followed phonotactics of English and Russian (native phonotactics) when they were stressed on the second (rather than first) syllable, e.g. *platok* [plaˈtok] than words with native phonotactics and word-initial stress, e.g. *pledik* [plʲedʲɪk]. Target words which followed phonotactics of English and Russian and which were stressed on the word-final syllable were the most accurately recognised words in both tasks, with the accuracy of 63%. In contrast, words which followed phonotactics of both English and Russian and stressed on the first syllable were recognised slightly above chance level in the word recognition task, and below chance (43%) in the forced-choice task. Moreover, words which followed phonotactics of only Russian (non-native phonotactics) and were stressed on the final syllable, e.g. *tvorec* [tvaˈrʲe͡ts] were recognised just slightly above chance level in both tasks, while words which followed phonotactics of only Russian and were stressed word-initially, e.g. *tvorog* [ˈtvorək] were recognised below chance level in the word recognition task and slightly above the chance level (52%) in the forced-choice task.

These results were unexpected, but they are very interesting and are discussed in the light what was already observed with respect to the effect of phonotactics and stress alone in Section 7.2.2. Firstly, it is not surprising that strong effect of native phonotactics is observed when phonotactics and stress cues interact because it was already discussed that effect of L1 native phonotactics was present in the word recognition task which was likely due to implicit knowledge being involved. Moreover, we previously observed marginal effect of word-final stress in the forced-choice task. It is likely that when these two cues (native phonotactics and strong-weak stress) interact, they reinforce each other and strengthen the effect. The effect of native phonotactics was predicted due to the L1 transfer and is not particularly surprising. However, the sensitivity to the weak-strong stress (typical in Russian) shows the ability to analyse a new language input as this stress pattern is not predominant in English. Moreover, the result of the word recognition task that words which followed only Russian phonotactics and were stressed on the final syllable were recognised below chance level can be explained on the basis of implicit knowledge, because words with Russian but not English phonotactics which are stressed on the final syllable are exactly what participants were not expected to transfer and had to acquire in the new language input. On the contrary, the result of the forced-choice task that words which followed English and Russian phonotactics and were stressed on the first syllable were recognised only at 43% could be due to forced-choice task being more of an explicit knowledge test, with participants invoking a conscious strategy of preferring words which do not "sound English".

The most interesting finding, however, is the fact that sensitivity to the combination of weak-strong stress and native phonotactics was remarkably robust, so it can be observed in both tasks, overcoming the implicit vs explicit knowledge effect discussed in the previous sub-sections. Since there was no significant effect of L2 Russian phonotactics even in the forced-choice (explicit knowledge) task, it is likely that sensitivity to L2 stress pattern is stronger than sensitivity to L2 phonotactics the knowledge of which cannot be based on L1 transfer.

Jusczyk *et al.* (1993a) showed that infants in monolingual English-speaking families could discriminate English from Norwegian but not from Dutch because in Norwegian (unlike English) pitch increases towards the end of the word, while Dutch and English have very similar prosodic patterns. They concluded that infants at six months of age already attended to the prosodic pattern of English, which is before they can attend to segmental and phonotactic information of English. In contrast, infants develop the sensitivity to language-specific phonotactics only by nine months (e.g. Friederici and Wessels 1993; Jusczyk *et al.* 1993a; Jusczyk *et al.* 1994; Mattys & Jusczyk 2001). Drawing the parallels between L1 acquisition by infants in those studies and L2 acquisition by adults in the present study, one can conjecture that adult *ab-initio* learners develop sensitivity to the prosodic pattern of a novel language before they develop sensitivity to new language-specific phonotactics. Carefully designed studies on *ab-initio* learners are needed to investigate this further.

7.2.3.2 Effect of combination of phonotactics and word length

The 6th research question of the present study asked whether *ab-initio* learners are sensitive to the interaction between phonotactics and word length when detecting words from the input. It was predicted that participants would be more accurate in detecting targets which followed phonotactics of both English and Russian, and they would be more accurate on bisyllabic than on monosyllabic words. However, the results of both tasks did not find a significant effect of *ab-initio* learners making use of interaction between phonotactics and word length. However, there was a trend in the word recognition task of bisyllabic words being recognised more accurately than monosyllabic words following both English and Russian phonotactics. Words following Russian but not English phonotactics were recognised much worse, although slightly above the chance level, and there was no preference for bisyllabic over monosyllabic words or on the other way round. These results are consistent with the explanation that *ab-initio* learners were drawing on implicit knowledge when performed on the word recognition task. In contrast to the word recognition task, there was no specific pattern in the forced-choice task, except for the fact that monosyllabic words with native phonotactics were recognised below chance level.

7.2.4 Effect of single cues over time

In this section, the results on the interaction of each cue (phonotactics, stress, word length) with the input are discussed. The 7th research question asked whether sensitivity to phonotactic constraints would increase over sessions. It was predicted that *ab-initio* learners would be more accurate in detecting words from the input with native phonotactics than non-native phonotactics, and this ability would increase over sessions. The 8th research question asked whether sensitivity to MSS (strong-weak stress pattern) would increase over sessions. It was predicted that *ab-initio* learners would be more accurate in detecting words which are stressed on the first syllable than on the last one, and this ability would increase over sessions. The 9th research question asked whether sensitivity to the word length would increase over sessions. It was predicted that *ab-initio* learners would be more accurate in detecting bisyllabic words than monosyllabic words and this ability would increase over sessions.

Surprisingly, given the robust effect of the number of sessions (increasing input) as described in Section 7.2.1, the results of both tasks showed that the effects of phonotactics, word stress and word length, when interacting with input, were not significant. Gullberg *et al.* (2012) found that interaction of frequency and bisyllables was a highly salient cue for Dutch participants to extract Chinese words after as little as seven minutes of exposure. In their study, words appeared either two times or eight times. In contrast, in the present study, the targets appeared only twice per session, which means that by the end of the fourth session participants encountered targets eight times, i.e. the same number of occurrences as in Gullberg *et al.* (2012) but with 56 rather than seven minutes of exposure. The absence of a significant effect of the interaction of bisyllabic words and frequency in the present study could be due to the different experimental setup or the languages.

Generally, the amount of exposure is known to have an important role in language acquisition (e.g. see N. Ellis 2003 for an overview on frequency effects in language processing). The present study demonstrated that the ability to detect targets in the word recognition task started to appear from the second session, and in the forced-choice task from the third session, see Section 7.2.1. It is consistent with Davis *et al.* (2009), who showed that words learnt the night before testing become consolidated in memory. Though the present study did not find significant effects of interaction between phonotactics, stress, and word length, with increasing amount of input, it is conceivable that with a larger sample such effects would appear.

7.2.5 Effect of generalisation

The 10th research question asked whether *ab-initio* learners can generalise to phonotactic properties heard in the input. It was predicted that in the word recognition task the ability to discriminate between targets and generalisable distractors (because they "sound similar" to targets) would be lower than that between targets and non-generalisable distractors. Similarly, in the forced-choice task, it was predicted that the false alarm rate on generalisable distractors would be similar to the hit rate on targets because the participants would easily confuse them with targets.

Both these predictions were supported by the experiments. The results from the word recognition task showed that *d'* index of sensitivity to generalisable distractors was indeed significantly lower than that on non-generalisable distractors. The results from the forced-choice task showed that participants indeed often confused generalisable distractors with targets – the false alarm rate for generalisable distractors was only slightly lower than the hit rate for targets.

The present study concerning generalisation demonstrated that *ab-initio* learners could generalise phonotactics of Russian to words which they did not encounter within the input when these novel words shared phonotactic properties with targets. There is only a handful of studies which investigated generalisation abilities in phonology at the early stages of language learning. The results of the present study with respect to generalisation are consistent with Gullberg *et al.* (2010) who found that L1 Dutch *ab-initio* learners could detect words from Mandarin and to generalise them to words they did not encounter in the input after as little as 7 minutes of exposure.

The present study also complements the results of Linzen and Gallagher (2017) who showed that adult English native speakers could rapidly generalise to new sounds after very little exposure to an artificial language. In their study they exposed participants to an artificial language where all words were of CVCV type, the onsets of which had the same voicing (i.e. they were either all voiced obstruents or all voiceless obstruents) while the rest of the word contained [l], [m], or [n] in the C-position with the stress on the first syllable. Similarly to the present study, the words were divided into three groups: (1) conforming attested onset, which meant that the word appeared in the input; (2) conforming novel onset, which meant that it did not appear in the input but had the same voicing as those which did; and (3) nonconforming unattested onset, which meant it was different in voice feature from the ones heard in the input. After the exposure phase participants were asked to participate in a task similar to the word recognition task in the present study. Their study utilised between-subject design where each participant was part of one of the four groups (one, two, four, or eight exposure sets). The results showed that after as little as one set of exposure participants could discriminate words with conforming attested onsets from words with nonconforming unattested onsets but participants started to differentiate words with conforming attested onsets from words with conforming unattested onsets only after two or more exposure sets.

The results of the present study, together with those of Gullberg *et al.* (2010) and Linzen and Gallagher (2017), demonstrate that generalisation can take place at the initial stages of both natural and artificial languages, i.e. the ability to make phonological generalisations is a fundamental property of language acquisition, and is observable after a very short exposure to a language.

7.3 Effect of linguistic training

The 11th research question asked whether there would be differences between linguistically sophisticated and linguistically naïve participants with respect to each of the research questions discussed above. Moreover, it was predicted that linguistically sophisticated participants would perform better than linguistically naïve participants in each case.

However, the present study showed that in general there was no significant difference in the performance between linguistically sophisticated and linguistically naïve participants, with the exception of the second hypothesis predicting that participants would be more accurate in detecting words following phonotactics of both English and Russian than words following only Russian phonotactics. In particular, the results of the forced-choice task were significant: Linguistically sophisticated participants correctly recognised 53% of targets with native phonotactics and 53% of targets with non-native phonotactics; in contrast, linguistically naïve participants were more accurate at detecting targets with non-native phonotactics (54%) while accuracy on words with native phonotactics was below chance at 47%. The results of the word recognition task were not significant, but there was a trend showing that linguistically sophisticated participants were more accurate than linguistically naïve ones. These results are interesting, and they fit into the explanation of the general findings from the word recognition and forced-choice task, see Section 7.2.2.1: In the word recognition task participants were drawing on implicit knowledge, while in the forced-choice task they were drawing on explicit knowledge (R. Ellis 2009). Consequently, the reason why there were no significant differences between linguistically sophisticated and linguistically naïve participants with respect to the effect of phonotactics in the word recognition task could be because this task drew on implicit knowledge, which is less susceptible to metalinguistic knowledge and the ability to analyse language consciously (e.g. Bialystok 1979; Elder *et al.* 1999; R. Ellis 2004).

On the contrary, in the forced-choice task, participants had a chance to demonstrate their explicit metalinguistic knowledge. Hence sophisticated participants could indeed rely on their metalinguistic knowledge, which made them not to show a preference for the non-native phonotactic patterns over the native ones, because certainly, both patterns exist in Russian. For instance, sophisticated participants could have known from their training in linguistics that if a language allows more complex structures (e.g. CC clusters with MSD=0 as in Russian), it should allow simpler structures (e.g. CC clusters with MSD=2 which are found in both English and Russian), see Eckman (1977). In contrast, linguistically naïve participants were also drawing on their explicit knowledge which made them avoid structures similar to those in their native language English and use a conscious strategy of preferring words which "sound least English" (Kellerman 1979). Since the naïve participants do not possess considerable metalinguistic knowledge (as none of them was trained in linguistics or had advanced knowledge of a foreign language), they were less likely to suppose that what exists in their native language (CC clusters with MSD=2) may exist in the target language.

It is surprising the present study did not find statistically significant differences between linguistically sophisticated and linguistically naïve participants for other research questions. However, there was a general trend of linguistically sophisticated participants to perform better than linguistically naïve with respect to almost all research questions. Hence it is conceivable that a larger sample could produce significant effects. (In the present study the linguistically sophisticated group comprised 15 participants, and the linguistically naïve group comprised 13 participants.) It would be interesting to investigate the same research questions on a larger sample.

7.4 Discussion of the cognate identification task

The main purpose of the cognate identification task was to check whether participants would be paying attention to the input while listening to it and to eliminate those participants who did not pay sufficient attention. The results on the cognate identification task showed that there was an improvement with cognates' detection from the first session to the second session, that is during that period when participants listened to the input with sentences containing cognates related to 'music'. Also, there was an improvement from the third session to the fourth session, or during that period, participants listened to input containing cognates related to 'university life'. However, there was no difference in the whole between participants' recognition of cognates related to 'music', and between recognition of cognate related to 'university life', but there was a trend that cognates related to the 'university life' were recognised better than cognates related to 'music' by the second time of exposure with each group. In other words, cognates recognition was better in the fourth session than in the second session. It is surprising, but it could be because participants got a gist of what they were tested on by the final testing on each group. Alternatively, it could because cognates related to university life are in general easier to detect since all participants were students. That is, the genre might have been familiar to them. It is conceivable, if a sample of musicians was tested on the same experiment, it could be that they would recognise better cognates related to 'music'. What is more, it is possible that none of these is a good explanation. Instead, the reason why participants' accuracy was higher on the cognates about 'university life' than on cognates about 'music' lies in fine-grained phonetic and phonologies properties of the target words. Carroll (2012) mentions that it is still not fully understood what is about phonetic and phonological properties of words which make them appear as similar or dissimilar enough.

In general, the results on the cognates are consistent with studies by Rast and Dommergues (2003) , Rast (2008), Rast (2010), and Shoemaker and Rast (2013), as well as Carroll (2014), who found that recognition of cognates or highly transparent words between L1 and L2 is robust even with as little as *no* input, or after a few trials of exposure to sentences containing cognate names. In fact, the results which were found in this task are perhaps even more striking, as it adds to the previous finding that learners can recognise cognates not only when they are presented with the isolated forms (Rast 2010; Shoemaker & Rast 2013) or in syntactic frames which varies in several ways as in Carroll (2014), but learners can also detect cognates when they are embedded in sequential contexts, for instance:

a.	*Igrat'*	*na*	***pianino***	*my*	*ychilis'*	*Davno*
	Play (INF)	on	piano (ACC)	we	learned	long time ago

We learned to play piano long time ago
[ɪgrˈatʲ nə pʲɪɐnʲˈinə mˈɨ ʊ͡tɕˈilʲɪsʲ davnˈo]

b.	*V*	*etom*	*gody*	***egzamen***	*budet*	*letom*
	In	this	year (PREP)	exam	(FUT)	summer (INSTR)

The exam will be in the summer this year
[v‿ˈɛtəm gˈodʊ ɪgzˈamʲɪn bˈudʲɪt lʲˈɵtəm]

As you see from the examples a. and b. above (see more examples in Appendix A1), there are no consistent phonotactic or prosodic cues which could have cued the recognition of cognates other than cognates themselves. With respect to the question of whether there were any participants in the present study who might not have been paying sufficient attention. No participants were identified whose performance was too low, so it did not fit with the pattern of overall responses. As a result, not a single participant was eliminated from the experiment based on a lack of attention.

7.5 Summary

Chapter 7 has provided a discussion of the results of the present study with respect to all research questions.

It was discussed that *ab-initio* learners' ability to detect target words of Russian increased with more input which they received on all four days. However, the accuracy effect varied depending on the task participants took. Specifically, the ability to detect Russian words increased in response to input on each day, with the difference being significant between all sessions except the second to third sessions on the word recognition task, but in the forced-choice task, the significant effect started to appear only from the third session. Substantial evidence suggests that L2 learners' native language determines which aspects of a target language can be acquired and which aspects are difficult to acquire. For instance, Eckman (1977) proposed that those aspects of the target language which are different and are more marked than those in the source language will be difficult. This was supported by different studies which looked at perception and production of acquisition of L2 phonotactics and stress (e.g. Broselow & Finer 1991; Archibald 1992, 1993; Carlisle 1991; Hart 1998; Ostapenko 2005). Moreover, such evidence of L1 transfer exists even in the highly proficient L2 learners whose L2 ability of L2 phonotactics was assessed using online psycholinguistic tasks (Altenberg & Cairns 1983; Weber 2000; Weber & Cutler 2006). It was discussed in this chapter that with respect to the L1 transfer of phonotactics, *ab-initio* learners of Russian were influenced by L1 phonotactic knowledge only in the word recognition task but not the forced-choice task. However, with respect to stress, *ab-initio* learners of Russian were not influenced by the L1 transfer of MSS for speech segmentation. Instead, they relied more on the opposite iambic stress, which is likely to be a default stress pattern in Russian as proposed by some researchers (e.g. Crosswhite *et al.* 2013). The sensitivity to the weak-strong stress pattern was evident only in the forced-choice task. The asymmetry of results between the word recognition and the forced-choice task was discussed in the light of implicit versus explicit memory processes which were likely to underpin the performances on each task (Hulstijn 2005; R. Ellis 2009). Moreover, these differences between the two tasks were likely the reason why the performance of the linguistically sophisticated group was not higher than that of linguistically naïve group except for when the effect of phonotactics was tested in the forced-choice task despite the significant results. In particular, linguistically sophisticated participants could exploit their metalinguistic knowledge in the forced-choice task, so their accuracy was above chance on words with native and non-native phonotactics. Naïve participants, however, could not exploit metalinguistic knowledge due to the absence of training in linguistics, but they seemed to make use of the strategy of accepting words which sounded least English. Interestingly enough, when phonotactics interacted with stress, regardless of the task, all participants were influenced most of all by native phonotactics and weak-strong stress pattern. It shows that the effect of native language transfer on phonotactics is strong, but the sensitivity to iambic stress pattern is likely to occur from *ab-initio* learners' attendance to the properties of Russian language as it is unlikely to be due to the L1 transfer because most of the polysyllabic words which exist in English spontaneous speech are stressed on the first syllable (Cutler & Carter 1987). What is more, English bisyllabic nouns are stressed word-initially 73% of all times as reported in Sereno (1986) and 94% of all times in Kelly and Block (1988), although it is highly unlikely that *ab-initio* learners of Russian were responding to the properties of the grammatical class.

Furthermore, it was discussed that *ab-initio* learners could generalise the phonotactics properties of Russian beyond what they heard in the Russian language input, and this ability was not affected by the type of task. Surprisingly the study did not find any effect of word length, and interaction between phonological cues and input, despite trends which pointed out that bisyllabic words were slightly preferred over monosyllabic words. Finally, it was discussed that the results on the cognate identification task conformed with the others on the initial stages of L1 learners that an ability to segment cognates after brief input is very powerful which was taken as an indication that all participants were listening to the input attentively and therefore engaging in the experiment as expected.

Chapter 8. Conclusions

8.1 Summary and conclusions

Infants with normal hearing ability learn how to convert continuous strings of sounds in their native languages into discrete meaningful units (the words of the language) before they learn to associate them with meanings and use them a number of meaningful ways, which is a formidable challenge. Extensive research on L1 speech perception shows that within the first year of life infants respond to the properties of many individual cues which may facilitate the identification of word boundaries, such as the sound patterns of their own names, and allophonic, phonotactic, prosodic and distributional cues, as well as the interaction of phonotactics with stress (e.g. Friederici & Wessels 1993; Jusczyk *et al.* 1993a, 1993b, 1994, 1999a,1999b; Hohne & Jusczyk 1994; Mandel *et al.* 1995; Saffran *et al.* 1996a).

This ability to rapidly analyse a continuous speech stream of an unknown language exists in adults who were exposed for the first time to an artificial language (e.g. Saffran *et al.* 1996b; Aslin *et al.* 1998; Chambers *et al.* 2003; Linzen & Gallagher 2017) or a natural language (e.g. Rast & Dommergues 2003; Rast 2008, 2010; Gullberg *et al.* 2010, 2012; Shoemaker & Rast 2013; Carroll 2012, 2014).

In this thesis, I have extended the findings of previous research on adults' segmentation abilities in a completely unfamiliar natural language by looking at a new language pair, and measuring English speaking adults' ability to segment Russian words from the input which was presented aurally for seven minutes on each day. I asked the following questions: (1) whether or not segmentation ability (the detection of words) would increase with increasing amounts of exposure over four consecutive days; (2) whether or not learners could rely on phonotactic, prosodic, and word length cues (as measured in terms of numbers of syllables) and if learners would be sensitive to the interactions between these cues; (3) whether or not sensitivity to these cues would increase over sessions; (4) whether or not learners would generalise beyond what they had heard in the input; and (5) whether or not linguistically sophisticated participants would perform better than linguistically naïve participants.

The participants in the present study were familiarised with target words in completely unfamiliar sequential contexts in Russian, and were then tested on the recognition of isolated versions of these words mixed with distractors. The main findings of the present study are as follows. The ability to recognise isolated forms of target words increased over four consecutive sessions. However, this effect of word recognition was more pronounced in the word recognition task than in the forced-choice task. These results are likely to be due to the fact the forced-choice task was more difficult than the word recognition task which made participants exploit explicit knowledge, while performance on the word recognition task was largely unconscious. These word detection abilities reflected the influence of L1 English phonotactic knowledge and sensitivity to weak-strong stress, as well as the interaction of these two cues, which is likely to stem from the analysis of novel language input. Besides, this study showed that *ab-initio* learners can generalise implicitly learned phonotactic information of words which they heard during the input to the novel examples which conformed to the phonotactics information they had heard during the input phase. Furthermore, the study showed a considerably robust effect of the identification of cognates which belonged to two groups of semantically unrelated words. This was taken to be an indication that participants were not fatigued to the extent of not paying sufficient attention to the input. Finally, this study is unique in examining linguistically sophisticated and linguistically naïve participants. The results of this examination showed that, in general, performance of linguistically sophisticated participants was not significnatly better than that of naïve participants.

Overall, the study suggests that adults have a mental capacity to identify isolated forms of words after being presented with these words embedded in sequences of speech without instruction. This ability is powerful, as it manifests itself in several other natural language pairs (see Gullberg *et al.* 2010, 2012; Carroll 2012, 2014; Shoemaker & Rast 2013). It appears that adults share this capacity with infants. This capacity is unconscious or implicit, as it is evident in the ability of infants and adults' to respond according to their intuition. However, there is evidence that adults could exploit explicit memory processes, as was observed during their performance in the forced-choice task in the present study.

One may argue that the segmentation abilities observed in the present study result from the participants' ability to listen to oral information, take it in, and recall it again after a short delay, known as *auditory memory* (Dawai and Cowan 2014). It is considered to be a part of general cognition and has nothing to do with language processing or parsing in a rationalist or "nativist" perspective on the representation and acquisition of linguistic knowledge. Chomsky is one of the most famous cognitive scientists who has worked under a rationalist approach. He is famous for his proposal that the source of linguistic knowledge is the mind rather than external input, which is the main argument of empiricists (e.g. Tomasello 2000a,b)

> "There is a specific faculty of the mind/brain that is responsible for the use and acquisition of language, a faculty with distinctive characteristics that is apparently unique to the species in essentials." (Chomsky 1987, 50).

The results of the present study with respect to generalisation abilities of phonotactic regularities showed that *ab-initio* learners could generalise phonotactics of Russian they encountered during the input to new words which they did not hear in the input. In other words, participants treated words not from the input as if those words they heard in the input when the phonotactic properties between those two groups of words were the same. On the contrary, participants knew that words phonotactic properties of which they did not hear in the input were not targets. These results illustrate that participants responses were not just based on their ability to listen to the target words during the input and then recall them at the testing phase; rather, the observed generalisation ability indicates that participants were constrained, systematic, and perhaps even creative as they went beyond the stimuli instead of mere copying or memorising of what they heard in the input.

The present study was inspired by previous studies in the *ab-initio* learners' paradigm, research into artificial languages, and studies of segmentation abilities by infants, and the methodology was mainly influenced by Gullberg *et al.* (2010, 2012). As already discussed, the accuracy rates in the present study were generally comparable to those reported in Gullberg *et al.* (2012). As already mentioned in Gullberg *et al.* (2012: 259), "the above chance performance is very different from successful L2 acquisition", but this ability, however modest, – represent the earliest steps in acquiring a new language. Moreover, it is crucial for our understanding of the L2 initial state (Vainikka & Young-Scholten 1994, 1996, 1998; Schwartz & Eubank 1996; Carroll 1999, 2001) and the theory of L2 acquisition in general. We do not know yet about what happens next during the actual acquisition of L2 until we conduct studies with foreign language learners who will have accumulated at least 10 months of exposure to an L2, akin to those of Peter W. Jusczyk and his colleagues.

8.2 Limitations and suggestions for future research

This section discusses the limitations of the present study and suggestions for future research.

The present study attempted to answer many research questions, and there was a need to manipulate several conditions such as (1) the amount of input to measure word detection ability in response to increasing input; (2) phonotactics; (3) stress; (4) word length to measure the response to cues as well as interactions between cues; (5) type of stimulus to measure generalisation; (6) linguistic group to measure if there was any advantage of linguistics training; and (7) cognates to measure attentiveness to the input. Consequently, there were perhaps too many variables for a single study, whereas the tasks themselves appear to constitute another variable as the word recognition task was an implicit memory task, while the forced-choice task was more of an explicit memory task. Future studies should try to avoid looking at the effects of so many variables in a single study. In particular, one would need to think very carefully about whether to utilize implicit or explicit memory tasks when testing for the effect of linguistics training along with the effects of other variables, because some of the results in the present study were difficult to interpret and it was not clear if they were due to differences in tasks or differences in groups. One could also try to investigate if the reaction times can provide any interesting information with respect to the same research questions – in fact, this is what I intend to explore next.

It would be really interesting to see if the same effects which were observed in the present study would hold with a larger sample. Recruiting participants to take part in a within-subject design which lasts over four consecutive days is an expensive procedure for both the researcher and participants. Modern technology allows the recruitment of very large samples of participants over the internet. For instance, Amazon Mechanical Turk (MTurk) is a website which contains features for data collection virtually, which has been gaining popularity over the recent years. Results obtained by MTurk have been shown to be similar to those which were collected in the laboratory (Crump *et al.* 2013).

The present study utilised two groups of sentences with cognates in order to test if participants would be able to detect the cognates. This was a way of establishing if participants were paying attention to the input. However, as a result of this procedure, a decline in responses in the third session was observed in the word recognition and cognate identification tasks. These tasks are tests of implicit memory so the decrease in responses was not surprising. Future studies utilising cognates should avoid switching cognate types halfway through.

The results of the present study show that the *ab-initio* learners relied on the weak-strong stress pattern for the detection of Russian words. This likely stems from these learners' capacity to analyse the new language input because it was predicted that learners would transfer their knowledge of MSS (Cutler & Norris 1988; Cutler 1990) when detecting words in Russian. Despite the fact that Crosswhile *et al.* (2003) has proposed that the Russian default stress pattern is iambic, which nicely explains why learners in the present study responded to the weak-strong pattern, other researchers disagree that Russian is iambic (e.g. Idsardi 1992; Halle 1997). Consequently, if we want to understand how learners respond to the prosodic pattern of a target language, the stress pattern of which differs from the L1, it would be sensible to select a target language for which the stress pattern is well understood.

Finally, as Carroll (2013) has previously observed, an important limitation of all studies in the *ab-initio* learners paradigm is that we still do not know how L2 learners at the beginning stages of L2 acquisition respond to variations in speaking rates and the talkers' voice which is typical of normal speech and which we know infants do very successfully within the first few months of life well before they start producing their first words (DeCasper & Fifer 1980; Eimas & Miller 1981; Kuhl 1985; Jusczyk *et al.* 1992). Future studies could address this gap in knowledge.

Appendix A.1 Input sentences

Target words are in bold, and sentences with cognates are underlined.

A1.1 Input sentences containing cognates about “music” played on the first and second day

Dobryj den’. Menia zovyt Nataliya!
dˈobrɨj dʲˈenʲ mʲɪnʲˈa zavˈut natˈalʲjə

Segodnia ya budu govorit’ o muzuki
sʲɪvˈodnʲə jˈæ bˈudu gəvarʲˈitʲ a mˈuzɪkʲɪ

1. On obiazan **knigam** za obrazovanie
ˈon abʲˈazən **knʲˈigəm** zə abrəzavˈanʲɪjə

2. Budet sostavlen **slovar’** na novyj god
bˈudʲɪt sastˈavlʲɪn **slavˈarʲ** nə nˈovɨj gˈot

3. Moj vysokij **shkaf** zabit bitkom
mˈoj vɨsˈokʲɪj **ʃkˈaf** zabʲˈit bʲˈitkom

4. Opiat doroshaet **hleb** v etom meste
apʲˈatʲ dəraʒˈajɪt **xlʲˈep** v ˈɛtəm mʲˈesʲtʲɪ

5. Muschina kushal **knel’** na veranda
muʃˀːˈinə kˈʊʃəl **knʲˈelʲ** nə vʲɪrˈanʲdʲɪ

6. <u>V etoj komante **akustika** otlichnaya</u>
<u>v ˈɛtəj kˈomnətʲɪ **akˈusʲtʲɪkə** atlʲˈitʃnəjə</u>

7. Chelovek videl **kluchok** sinevy
tʃɪlavʲˈek vʲˈidʲɪl **klatʃok** sʲɪnʲɪvˈɨ

8. Odnako poluchil **srostok** cherez den’
adnˈakə pəlʊtʃˀˈil **srˈostək** tʃˀerʲɪz dʲˈenʲ

9. Mal’chik videl **grom** na golubom nebe
mˈalʲtʃɪk vʲˈidʲɪl **grˈom** nə gəlʊbˈom nʲˈebʲˈe

10. V otlichii ot basa **bariton** imeet perehodnyj ton
v atlʲˈitʃʼɪɪ ad bˈasə, **bərʲɪtˈon** ɪmʲˈejɪt pʲɪrʲɪxˈodnɨj tˈon

11. Unosha byl **grychik** horoshij
jˈunəʃə bˈɨl **grˈuʃʼːɪk** xarˈoʃɨj

12. On lubil **zvyk** budil’nika
ˈon lʲubʲˈil **zvˈuk** bʊdʲˈilʲnʲɪkə

13. Ne veril **smenschik** begal dolgo
nʲˈe vʲˈerʲɪl **smʲˈenʲʃʼːɪk** bʲˈegəl dˈolgə

14. Nam nuzhen **pledik** pod divan
nˈam nˈuʒˈɨn **ˈpledik** pəd dʲɪvˈan

15. Tam igrala **gitara** i **sintezator** nailuchshego kachestva
tˈam ɪgrˈalə gʲɪtˈarə ˈi sʲɪnʲtʲɪzˈatər nəɪlˈutʃʼʃɨvə kˈatʃʼɪstvə

16. Ya ne lublu kushat’ **blinnik** na obed
jˈæ nʲˈe lʲublʲˈu kˈuʃətʲ **blʲˈinʲːɪk** nə abʲˈet

17. Tot zolotoj **slitok** syschestvoval davno
tˈot zəlatˈoj **slʲˈitək** sʊʃʼːɪstvavˈal dˈavno

18. Yzhe nastupil **srok** peremiriya
ʊʒɛ nəstʊpʲˈil **srˈok** pʲɪrʲɪmʲˈirʲɪjə

19. Paren’ govoril **zvonar’** rabotal nedelu
pˈarʲɪnʲ gəvarʲˈil **zvanˈarʲ** rabˈotəl nʲɪdʲˈelʲu

20. My slyshali **gimn** po radio
mˈɨ slˈuʃəlʲɪ **gʲˈimn** pə rˈadʲɪo

21. Oni lubyat **smes’** finikov i orehov
anʲˈi lʲˈubʲɪt **smʲˈesʲ** fʲˈinʲɪkəf ˈi arʲˈexof

22. Horoshij **kniazek** pomog im
xarˈoʃɨj **knʲɪzʲˈok** pamˈok ˈim

23.	Ochen’	milyj	**krolik**	pil
	ˈotʃɪnʲ	mʲˈilɨj	**krˈolʲɪk**	pʲˈil

24.	Novyj	**shkol’nik**	ponimaet	zadachy
	nˈovɨj	**ʃkˈolʲnʲɪk**	pənʲɪmˈajɪt	zadˈatʃˈu

25.	Na	vechere	**ekspoziciya**	byla	garmonichnoj
	nə	vʲˈetʃˈɪrʲɪ	**ɪkspazʲˈitsɨjə**	bɨlˈa	gərmanʲˈitʃˈnəj

26.	Kucheryavyj	**hlopec**	zahodil	Na	chaj
	kʊtʃˈɪrʲˈævɨj	**xlˈopʲɪts**	zəxadʲˈil	nə	tʃˈæj

27.	Y	rebyat	dorogoj	**tvid**	Lubimyj	naryad
	ʊ	rʲɪbʲˈæt	darˈogəj	**tvʲˈit**	lʲubʲˈimɨj	narʲˈæt

28.	My	lubim	**tvorog**	s	Izumom
	mˈɨ	lʲˈubʲɪm	**tvarˈok**	s	ɪzʲˈuməm

29.	Bez	kolebanij	**shtopar**	ne	budet	lishnij
	bʲɪs	kəlʲɪbˈanʲɪj	**ʃtˈopər**	nʲˈe	bˈudʲɪt	lʲˈiʃnʲɪj

30.	Altovyj	**saksophon**	eto	Duhovnyj	myzykal’nyj	instrument
	alʲtˈovɨj	**səksafˈ**	ˈɛtə	dʊxˈovnɨj	mʊzɨkˈalʲnɨj	ɪnstrʊmʲˈent

31.	Sud’ya	dal	**svistok**	potom	nachalas’	p’esa
	sʊˈdʲja	dˈal	**svʲɪstˈok**	pˈotəm	nətʃˈɪlˈasʲ	ˈpˈjesa

32.	Horoshij	**sluh**	ne	obhodim	Nam
	xarˈoʃɨj	**slˈux**	nʲˈe	apxˈodʲɪm	nˈam

33.	Moj	sinij	**platok**	lezhit	na	poly
	mˈoj	sʲˈinʲɪj	**plɐtˈok**	lʲɪaʒˈɨt	Na	pˈolʊ

34.	Inogda	peredat’	**shtyrval**	Mozhno
	ɪnagdˈa	pʲɪrʲɪdˈatʲ	**ʃtʊrvˈal**	mˈoʒnə

35.	Ih	duet	hot’	natyral’nay	no	**liriky**	nikto	ne	ponimaet
	ˈix	dʊˈɛt	xˈotʲ	nətʊrˈalʲnɨj	nˈo	**lʲˈirʲɪkʊ**	nʲɪktˈo	nʲˈe	pənʲɪmˈajɪt

36.	S	ego	zhenoj	**smugliak**	zhil	ladno
	s	jɪvˈo	ʒɨnˈoj	**smuglʲˈæk**	ʒˈɨl	lˈadnə

37. Kak delaet **shkodnik** tak ne delaut
kˈak dʲˈeləjɪt **ʃkˈodnʲɪk** tˈak nʲˈe dʲˈeləjut

38. V etom **bliznec** pohodil na nas
v ˈɛtəm **blʲɪzʲnʲˈets** pəxadʲˈil nə nˈas

39. Dazhe legkij **blef** nel’zya Najti
dˈaʒɨ lʲˈoxkʲɪj **blʲˈef** nʲɪlʲzʲˈæ najtʲˈi

40. Eto byla **muzuka** simphonicheskogo orkestra i hora
ˈɛtə bɨlˈa **mˈuzɨkə** sʲɪmfanʲˈitʃˈɪskəvə arkʲˈestrə ˈi xˈorə

41. Lubitel’ kupil **graphin** po deshevoj cene
lʲubʲˈitʲɪlʲ kʊpʲˈil **grafʲˈin** pə dʲɪʃˈovəj tsˈɛnə

42. Eti rabochie sushat **klever** na solnce
ˈɛtʲɪ rabˈotʃˈɪjə sˈuʃət **klʲˈevʲɪr** nə sˈontsɨ

43. Nedorogoj **zvonok** s etogo operatora
nʲɪdərag'oj **zvanˈok** s ˈɛtəvə apaʲɪˈratərə

44. Nezhalatel’nyj **krah** minuvshego
nʲɪʒɨlˈatʲɪlʲnɨj **krˈax** mʲɪnˈufʃɨvə

45. Igrat’ na **pianino** my ychilis’ davno
ɪgrˈatʲ nə **pʲɪanʲˈino** mˈɨ utʃˈˈilʲɪsʲ davnˈo

46. Esli on zhelaet **plov** delaite Iz baraniny
jˈeslʲɪ ˈon ʒɨlˈajɪt **plˈof** dʲˈeləjtʲɪ ɪz barˈanʲɪnɨ

47. Ego zolotoj **klik** perelivalsia
jɪvˈo zəlatˈoj **klˈɨk** pʲɪrʲɪlʲɪvˈalsʲə

48. Sovetuut kupit’ **krovat’** luboj shiriny
savʲˈetʊjut kʊpʲˈitʲ **kravˈatʲ** lʲubˈoj ʃɨrʲɪnˈɨ

49. Rumiannyj **glianec** byl na eyo schekah
rʊˈmʲanɨj **glʲˈanʲɪts** bˈɨl nə jɪjˈo ʃˈːɪkˈax

50.	**Soprano**	eto	vysokij	zhenskij	golos			
	saprˈanə	ˈɛtə	vɨsˈokʲɪj	ʒˈɛnskʲɪj	gˈoləs			
51.	Sosed	hotel	**glatok**	goriachej	vody			
	sasʲˈet	xatʲˈel	**glaˈtok**	garʲˈætʃˈɪj	vadˈɨ			
52.	Rabochij	sobral	**svitok**	zhivo				
	rabˈotʃˈɪj	sabrˈal	**svʲˈitək**	ʒˈɨvə				
53.	Eta	**melodiya**	byla	ochen’	horoshej			
	ˈɛtə	**mʲɪlˈodʲɪjə**	bɨlˈa	ˈotʃˈɪnʲ	xarˈoʃɨj			
54.	V	lesy	hodil	**srybschik**	dereviev			
	v	lʲɪsˈu	xadʲˈil	**srˈupʃˈːɪk**	dʲɪrʲˈevʲjɪf			
55.	Nikogda	ne	ponimal	**glaz**	Toj	devushki		
	nʲɪkagdˈa	nʲˈe	pənʲɪmˈal	**glˈas**	tˈoj	dʲˈevʊʃkʲɪ		
56.	Tihij	**hlopok**	donesjia	do	Neyo			
	tʲˈixʲɪj	**xlapˈok**	danʲˈosʲːə	də	nʲɪjˈo			
57.	Unosha	ponimaet	**tvorec**	schastiya	sam			
	jˈunəʃə	pənʲɪmˈajɪt	**tvarʲˈets**	sʃˈːˈasʲtʲjə	sˈam			
58.	Belyj	**svet**	razdelyaet	ih				
	bʲˈelɨj	**svʲˈet**	rəzʲdʲɪlʲˈæjɪt	ix				
59.	Nam	nuzhen	**shtat**	luchshe	Chen	u	teh	rebiat
	nˈam	nˈuʒɨn	**ʃtˈat**	lˈʊtʃˈʃɨ	tʃˈem	u	tʲˈex	rʲɪbʲˈat

A1.2 Input sentences containing cognates about "university life" played on the first and second days

Dobryj	den'.	Menia	zovyt	Natalia.	
dˈobrɨj	dʲˈenʲ	mʲɪnʲˈa	zavˈʊt	natˈalʲjə	

Segosnia	ya	budy	govorit'	o	zhisni	studentov
sʲɪvˈodnʲə	jˈæ	bˈudʊ	gəvarʲˈitʲ	o	ʒˈɨzʲnʲɪ	stʊdʲˈentəf

1.	Serij	tvid	visel	na	veshalke		
	sʲˈerɨj	tvʲˈit	vʲˈesʲɪl	nə	vʲˈeʃəlkʲɪ		
2.	Novey	smeschik	po	zadaniu	poyavilsya		
	nˈovɨj	smʲˈenʲʃːɪk	pə	zadˈanʲɪju	pəjɪvʲˈilsʲə		
3.	Vyduschij	geroj	krolik	v	etom	gody	
	vʲɪdˈuʃːɪj	gʲɪrˈoj	krˈolʲɪk	v	ˈɛtəm	gˈodʊ	
4.	Menia	ydivil	zvonok	v	veterinarnyu		
	mʲɪnʲˈa	ʊdʲɪvʲˈil	zvanˈok	v	vʲɪtʲɪrʲɪˈnarnʊu		
5.	Zainteresovannyj	**student**	pjet	Sok			
	zəɪnʲtʲɪrʲɪsˈovənːɨj	**stʊdʲˈent**	pʲjˈot	sˈok			
6.	Dorogoj	**slitok**	budet	podarkom			
	dəragˈoj	**slʲˈitək**	bˈudʲɪt	padˈarkəm			
7.	Eyo	verhnij	**klik**	bil	belyj		
	jɪjˈo	vʲˈerxnʲɪj	**klˈɨk**	bˈɨl	bʲˈelɨj		
8.	Annotacii	**knigam**	mi	napisali			
	anːatˈatsɨj	**knʲˈigəm**	mˈɨ	nəpʲɪsˈalʲɪ			
9.	On	voshol	kak	**shtopar**	v	ego	zadanie
	ˈon	vaʃˈol	kˈak	**ʃtˈopər**	v	jɪvˈo	zadˈanʲɪjə
10.	Etot	**profesor**	ego	partner			
	ˈɛtət	**prafʲˈesər**	jɪvˈo	partnʲˈor			

11. Lakomij **plov** my eli dosyta
lˈakəmɨj **plˈof** mˈɨ jˈelʲɪ dˈosɨtə

12. Yhod za kozhej **glaz** nuzhen lubomy
ʊxˈot zə kˈoʒɨj **glˈas** nˈuʒɨn lʲubˈomʊ

13. Vidimym **svet** delaet mir
vʲˈidʲɪmɨm **svʲˈet** dʲˈeləjɪt mʲˈir

14. Mozhet byt’ **zvonar’** zhenschina tozhe
mˈoʒɨt bˈɨtʲ **zvanˈarʲ** ʒˈɛnʲʃːɪnə tˈoʒɨ

15. <u>Zatem</u> <u>poluchit</u> <u>kvalifikaciu</u> <u>**doktor**</u> <u>nauk</u>
<u>zatʲˈem</u> <u>palˈutʃˈɪt</u> <u>kvəlʲɪfʲɪkˈatsɨju</u> <u>**dˈoktər**</u> <u>naˈuk</u>

16. Siyauschij **glianec** vyglyadel dorogo
sʲɪjˈajuʃːɪj **glʲˈanʲɪts** vˈɨglʲɪdʲɪl dˈorəgə

17. Nuzhno vylit’ **smes’** v Sotejnik
nˈuʒnə vˈɨlʲɪtʲ **smʲˈesʲ** f satʲˈejnʲɪk

18. Retivyj **groschik** pomogal Horosho
rʲɪtʲˈivɨj **grˈuʃʲːɪk** pəmagˈal xəraʃˈo

19. U nego etot **zvyk** poluchilsya
ʊ nʲɪvˈo ˈɛtət **zvˈuk** pəlʊtʃʲˈilsʲə

20. <u>Gosudarstvenyj</u> <u>**yniversitet**</u> <u>nahoditsia</u> <u>v</u> <u>centre</u> <u>goroda</u>
<u>gəsʊdˈarstvʲɪnːɨj</u> <u>**ʊnʲɪvʲɪrsʲɪtʲˈet**</u> <u>naxˈodʲɪtsə</u> <u>f</u> <u>tsˈɛntrʲɪ</u> <u>gˈorədə</u>

21. Medovyj **tvorok** nasha Lubimaya eda
mʲɪdˈovɨj **tvarˈok** nˈaʃə lʲubʲˈiməjə jɪdˈa

22. On pochyvstvoval **blef** v ih namerenoyah
ˈon patʃˈustvəvəl **blʲˈef** v ˈɨx namʲˈerʲɪnʲɪjɪx

23. Poleznyj **klever** lechit ot nedugov
palʲˈeznɨj **klʲˈevʲɪr** lʲˈetʃʲɪt at nʲɪdˈugəf

24. Paren’ kupil **krovat’** na toj nedeli

pˈarʲɪnʲ kʊpʲˈil **kravˈatʲ** nə tˈoj nʲɪdʲˈelʲɪ

25. Dervenskij **hlopec** ne lubil iskustva
dʲɪrʲɪvʲˈenskʲɪj **xlˈopʲɪts** nʲˈe lʲubʲˈil ɪskˈustvə

26. Budet zaputannyj **material** na poslednem kurse
bˈudʲɪt zapˈutənːɨj **mətʲɪrʲɪˈal** nə paslʲˈednʲɪm kˈursʲɪ

27. Torgovec prines **shkaf** cherez perehod
targˈovʲɪts pʲɪrʲɪnʲˈos **ʃkˈaf** tʃʲˈerʲɪs pʲɪrʲɪxˈot

28. Lish odin **glatok** moloka
lʲˈiʃ adʲˈin **glaˈtok** məlakˈa

29. Ee bydil **hlapok** v vosim chasov
jɪjˈo bʊdʲˈil **xlapˈok** f ˈvosʲɪmʲ tʃʲɪsˈof

30. Samaya interesnaya **lekciya** po sredam
sˈaməjə ɪnʲtʲɪrʲˈesnəjə **lʲˈektsɨjə** pə srʲˈedəm

31. Nam nuzhen **shtat** vernee bivshego
nˈam nˈuʒɨn **ʃtˈat** vʲɪrnʲˈejə bˈɨfʃɨvə

32. Samorodnyj **svinec** soderzhit piat’ izotopov
səmarˈodnɨj **svʲɪnʲˈets** sadʲˈerʒɨt pʲˈætʲ ɪzatˈopəf

33. Nuzhen **slovar’** vyrazhenij nam
nˈuʒɨn **slavˈarʲ** vɨraʒˈɛnʲɪj nam

34. Takii nayki kak **matematika** I **lingvistika** pohozhi
takʲˈijə naˈukʲɪ kˈak **mətʲɪmˈatʲɪkə** ˈi **lʲɪngvʲˈistkə** paxˈoʒɨ

35. On poterpel **krah** v itoge neudach
ˈon pətʲɪrpʲˈel **krˈax** v ɨtˈogʲɪ nʲɪʊdˈatʃʲ

36. Pohydel **smuglyak** posle Kanikyl
pəxʊdʲˈel **smʊˈglʲak** pˈoslʲɪ kanʲˈikʊl

37. Eto byl yzhasauschij **grom** And golovoj
ˈɛtə bˈɨl ʊʒasˈajuʃʲːɪj **grˈom** nəd gəlavˈoj

38. Eto pridymal **shkodnik** izvestnij
ˈɛtə prʲɪdˈuməl **ʃkˈodnʲɪk** ɪzvʲˈesnɨj

39. Ysidschivyj **shkol’nik** sidit na yrokah
ʊsʲˈitʃʲːɪvɨj **ʃkˈolʲnʲɪk** sʲɪdʲˈit nə ʊrˈokəx

40. V osnove idei **sistema** lizhit
v asnˈovʲɪ ɪdʲˈeɪ **sʲɪsʲtʲˈemə** lʲɪʒˈɨt

41. Zavershim **srok** v etom Iule
zəvʲɪrʃˈɨm **srˈok** v ˈɛtəm ɪjˈulʲɪ

42. Milyj **bliznec** pohodil na diadu
mʲˈilɨj **blʲɪzʲnʲˈets** pəxadʲˈil nə dʲˈædʲu

43. Emy ne dadut **shtyrval** poka on molod
jɪmˈu nʲˈe dadˈut **ʃtʊrvˈal** pakˈa ˈon mˈolət

44. Malen’kij **svitok** byl Korichnevyj
mˈalʲɪnʲkʲɪj **svʲˈitək** bˈɨl karʲˈitʃʲnʲɪvɨj

45. Yniversitetskaya **biblioteka** nedaleko ot kampusa
ʊnʲɪvʲɪrsʲɪtʲˈetskəjə **bʲɪblʲɪatʲˈekə** nʲɪdˈalʲɪkə ˈot kˈampʊsa

46. Tot luboj **srostok** raznyh vidov
tˈot lʲubˈoj **srˈostək** rˈaznɨx vʲˈidəf

47. Nam nuzhen **hleb** na yzhin
nˈam nˈuʒɨn **xlʲˈep** nə ˈuʒɨn

48. Ego nazyvaut **tvorec** Pirozhenyh
jɪvˈo nəzɨvˈajut **tvarʲˈets** pʲɪrˈoʒnɨx

49. Rodovityj **knizyok** voshel v Ih semiu
rədavʲˈitɨj **knʲɪzʲˈok** vaʃˈol v ˈɨx sʲɪmʲjˈu

50. Akademicheskij **simestr** zakanchivaetsya v mae
akədʲɪmʲˈitʃʲɪskʲɪj **sʲɪmʲˈestr** zakˈanʲtʃʲɪvəjɪtsə v mˈajə

51. Kto-to pystil **sluh** po gorodu
ktˈo-tə pʊsʲtʲˈil **slˈux** pə gˈorədʊ

52. Nuzhen **blinnik** kak ygoschenie
nˈuʒɨn **blʲˈinʲːɪk** kˈak ʊgaʃːˈenʲɪjə

53. Ogromij **klachok** bumagi lezhal
agrˈomnɨj **klaˈtʃʲ ok** bʊmˈagʲɪ lʲɪʒˈal

54. Tyt podaut **knel** novogo povara
tˈut pədajˈut **knʲˈelʲ** nˈovəvə pˈovərə

55. V etom gody **egzamen** budet letom
v ˈɛtəm gˈodʊ **ɪgzˈamʲɪn** bˈudʲɪt lʲˈotəm

56. Sosedi hotiat **pledik** pod Divan
sasʲˈedʲɪ xatʲˈæt **plʲedʲɪk** pəd dʲɪvˈan

57. Nevesomyj **platok** nodela ona
nʲɪvʲɪsˈomɨj **platˈok** nadʲˈelə anˈa

58. Schatlivyj **srybschik** rabotaet v sadu
ʃːɪslʲˈivɨj **srˈupʃːɪk** rabˈotəjɪt f sˈadʊ

59. Zelenyj **grafin** podarili im
zʲɪlʲˈonɨj **grafʲˈin** pədarʲˈilʲɪ ˈim

Appendix A.2 Experimental design for the word recognition task

	Word	Keyboard _response	Type	Phonotactics	Stress_length	Correct_ response
1	120	m	distractor	native	weak_strong	0
2	117	m	distractor	native	weak_strong	0
3	98	m	distractor	non_native	strong_weak	0
4	5	z	target	native	strong_weak	1
5	133	m	distractor	native	mono	0
6	94	m	distractor	non_native	mono	0
7	70	z	target	non_native	mono	1
8	76	z	target	non_native	mono	1
9	78	z	target	non_native	weak_strong	1
10	73	z	target	non_native	mono	1
11	77	z	target	non_native	strong_weak	1
12	134	m	distractor	native	strong_weak	0
13	6	z	target	native	weak_strong	1
14	111	m	distractor	native	weak_strong	0
15	132	m	distractor	native	weak_strong	0
16	56	z	target	non_native	strong_weak	1
17	185	m	distractor	neither	strong_weak	0
18	184	m	distractor	neither	mono	0
19	135	m	distractor	native	weak_strong	0
20	102	m	distractor	non_native	weak_strong	0
21	160	m	distractor	neither	mono	0
22	75	z	target	non_native	weak_strong	1
23	104	m	distractor	non_native	strong_weak	0
24	18	z	target	native	weak_strong	1
25	45	m	distractor	native	weak_strong	0
26	176	m	distractor	neither	strong_weak	0
27	35	m	distractor	native	strong_weak	0
28	65	z	target	non_native	strong_weak	1
29	12	z	target	native	weak_strong	1
30	84	m	distractor	non_native	weak_strong	0
31	115	m	distractor	native	mono	0
32	166	m	distractor	neither	mono	0
33	175	m	distractor	neither	mono	0
34	19	z	target	native	mono	1
35	59	z	target	non_native	strong_weak	1
36	128	m	distractor	native	strong_weak	0
37	93	m	distractor	non_native	weak_strong	0

38	99	m	distractor	non_native	weak_strong	0
39	37	m	distractor	native	mono	0
40	13	z	target	native	mono	1
41	24	z	target	native	weak_strong	1
42	164	m	distractor	neither	strong_weak	0
43	23	z	target	native	strong_weak	1
44	67	z	target	non_native	mono	1
45	32	m	distractor	native	strong_weak	0
46	22	z	target	native	mono	1
47	165	m	distractor	neither	weak_strong	0
48	4	z	target	native	mono	1
49	116	m	distractor	native	strong_weak	0
50	90	m	distractor	non_native	weak_strong	0
51	44	m	distractor	native	strong_weak	0
52	50	m	distractor	native	strong_weak	0
53	161	m	distractor	neither	strong_weak	0
54	72	z	target	non_native	weak_strong	1
55	8	z	target	native	strong_weak	1
56	33	m	distractor	native	weak_strong	0
57	95	m	distractor	non_native	strong_weak	0
58	129	m	distractor	native	weak_strong	0
59	57	z	target	non_native	weak_strong	1
60	167	m	distractor	neither	strong_weak	0
61	48	m	distractor	native	weak_strong	0
62	55	z	target	non_native	mono	1
63	110	m	distractor	native	strong_weak	0
64	58	z	target	non_native	mono	1
65	119	m	distractor	native	strong_weak	0
66	113	m	distractor	native	strong_weak	0
67	88	m	distractor	non_native	mono	0
68	122	m	distractor	native	strong_weak	0
69	168	m	distractor	neither	weak_strong	0
70	69	z	target	non_native	weak_strong	1
71	100	m	distractor	non_native	mono	0
72	74	z	target	non_native	strong_weak	1
73	131	m	distractor	native	strong_weak	0
74	163	m	distractor	neither	mono	0
75	186	m	distractor	neither	weak_strong	0
76	91	m	distractor	non_native	mono	0
77	130	m	distractor	native	mono	0
78	47	m	distractor	native	strong_weak	0
79	29	m	distractor	native	strong_weak	0

80	38	m	distractor	native	strong_weak	0
81	14	z	target	native	strong_weak	1
82	121	m	distractor	native	mono	0
83	43	m	distractor	native	mono	0
84	40	m	distractor	native	mono	0
85	28	m	distractor	native	mono	0
86	20	z	target	native	strong_weak	1
87	30	m	distractor	native	weak_strong	0
88	183	m	distractor	neither	weak_strong	0
89	123	m	distractor	native	weak_strong	0
90	66	z	target	non_native	weak_strong	1
91	11	z	target	native	strong_weak	1
92	82	m	distractor	non_native	mono	0
93	41	m	distractor	native	strong_weak	0
94	51	m	distractor	native	weak_strong	0
95	169	m	distractor	neither	mono	0
96	86	m	distractor	non_native	strong_weak	0
97	31	m	distractor	native	mono	0
98	101	m	distractor	non_native	strong_weak	0
99	34	m	distractor	native	mono	0
100	17	z	target	native	strong_weak	1
101	109	m	distractor	native	mono	0
102	182	m	distractor	neither	strong_weak	0
103	21	z	target	native	weak_strong	1
104	36	m	distractor	native	weak_strong	0
105	87	m	distractor	non_native	weak_strong	0
106	10	z	target	native	mono	1
107	89	m	distractor	non_native	strong_weak	0
108	9	z	target	native	weak_strong	1
109	179	m	distractor	neither	strong_weak	0
110	171	m	distractor	neither	weak_strong	0
111	180	m	distractor	neither	weak_strong	0
112	2	z	target	native	strong_weak	1
113	97	m	distractor	non_native	mono	0
114	3	z	target	native	weak_strong	1
115	42	m	distractor	native	weak_strong	0
116	64	z	target	non_native	mono	1
117	62	z	target	non_native	strong_weak	1
118	170	m	distractor	neither	strong_weak	0
119	49	m	distractor	native	mono	0
120	16	z	target	native	mono	1
121	114	m	distractor	native	weak_strong	0

122	39	m	distractor	native	weak_strong	0
123	92	m	distractor	non_native	strong_weak	0
124	181	m	distractor	neither	mono	0
125	103	m	distractor	non_native	mono	0
126	1	z	target	native	mono	1
127	68	z	target	non_native	strong_weak	1
128	63	z	target	non_native	weak_strong	1
129	7	z	target	native	mono	1
130	61	z	target	non_native	mono	1
131	162	m	distractor	neither	weak_strong	0
132	178	m	distractor	neither	mono	0
133	118	m	distractor	native	mono	0
134	85	m	distractor	non_native	mono	0
135	60	z	target	non_native	weak_strong	1
136	96	m	distractor	non_native	weak_strong	0
137	46	m	distractor	native	mono	0
138	127	m	distractor	native	mono	0
139	112	m	distractor	native	mono	0
140	71	z	target	non_native	strong_weak	1
141	105	m	distractor	non_native	weak_strong	0
142	177	m	distractor	neither	weak_strong	0
143	15	z	target	native	weak_strong	1
144	83	m	distractor	non_native	strong_weak	0

Appendix A.3 Experimental design from the forced-choice task

	Word1	keyboard_ response	Type	Phonotactics	Stress_length	Correct _response1	Word2
1	102	m	distractor	non_native	weak_strong	0	75
2	63	m	target	non_native	weak_strong	1	90
3	2	z	target	native	strong_weak	1	29
4	10	z	target	native	mono	1	37
5	78	m	target	non_native	weak_strong	1	105
6	12	z	target	native	weak_strong	1	39
7	100	m	distractor	non_native	mono	0	73
8	83	m	distractor	non_native	strong_weak	0	56
9	57	m	target	non_native	weak_strong	1	84
10	82	m	distractor	non_native	mono	0	55
11	60	m	target	non_native	weak_strong	1	87
12	43	z	distractor	native	mono	0	16
13	104	m	distractor	non_native	strong_weak	0	77
14	66	m	target	non_native	weak_strong	1	93
15	13	z	target	native	mono	1	40
16	30	z	distractor	native	weak_strong	0	3
17	41	z	distractor	native	strong_weak	0	14
18	65	m	target	non_native	strong_weak	1	92
19	33	z	distractor	native	weak_strong	0	6
20	76	m	target	non_native	mono	1	103
21	74	m	target	non_native	strong_weak	1	101
22	62	m	target	non_native	strong_weak	1	89
23	9	z	target	native	weak_strong	1	36
24	31	z	distractor	native	mono	0	4
25	38	z	distractor	native	strong_weak	0	11
26	17	z	target	native	strong_weak	1	44
27	97	m	distractor	non_native	mono	0	70
28	72	m	target	non_native	weak_strong	1	99
29	46	z	distractor	native	mono	0	19
30	47	z	distractor	native	strong_weak	0	20
31	69	m	target	non_native	weak_strong	1	96
32	49	z	distractor	native	mono	0	22
33	32	z	distractor	native	strong_weak	0	5
34	58	m	target	non_native	mono	1	85
35	48	z	distractor	native	weak_strong	0	21
36	98	m	distractor	non_native	strong_weak	0	71
37	18	z	target	native	weak_strong	1	45

38	34	z	distractor	native	mono	0	7
39	91	m	distractor	non_native	mono	0	64
40	23	z	target	native	strong_weak	1	50
41	59	m	target	non_native	strong_weak	1	86
42	15	z	target	native	weak_strong	1	42
43	61	m	target	non_native	mono	1	88
44	95	m	distractor	non_native	strong_weak	0	68
45	67	m	target	non_native	mono	1	94
46	1	z	target	native	mono	1	28
47	8	z	target	native	strong_weak	1	35
48	51	z	distractor	native	weak_strong	0	24

Appendix A.4 Experimental design for the cognate identification task

	Word	Keyboard_response	Correct_response
1	1	z	1
2	11	m	0
3	2	z	1
4	12	m	0
5	13	m	0
6	14	m	0
7	15	m	0
8	3	z	1
9	4	z	1
10	5	z	1
11	16	m	0
12	6	z	1
13	17	m	0
14	7	z	1
15	8	z	1
16	9	z	1
17	10	z	1
18	18	m	0
19	20	m	0
20	19	m	0

Appendix A.5 Lists of cognates and foils

About 'music'

	Target Cognate			Distractor		
	Russian	Transcription	Translation	Russian	Transcription	Translation
1	akustika	[aˈkusʲtʲɪkə]	acoustics	bunet	[brʲʉˈnʲet]	brunette
2	bariton	[bərʲɪˈton]	baritone	budjet	[bʲʊdˈʒɛt]	budget
3	gitara	[gʲɪˈtarə]	guitar	director	[dʲɪˈrʲektər]	director
4	soprano	[saˈpranə]	soprano	komjuter	[kam(ʲ)ˈp(ʲ)jutɨr]	computer
5	duet	[dʊˈɛt]	duet	nomer	[ˈnomʲɪr]	number
6	lirika	[ˈlʲirʲɪkə]	lyric	robot	[ˈrobət]	robot
7	myzyka	[ˈmuzɨkə]	music	telefon	[tʲɪlʲɪˈfon]	telephone
8	pianino	[pʲɪaˈnʲinə]	piano	televizor	[tʲilʲiˈvʲizar]	television
9	sintezator	[sʲɪntɨˈzatər]	synthesizer	hokej	[xaˈkʲej]	hockey
10	milodiya	[mʲɪˈlodʲɪjə]	melody	futbol	[fʊdˈbol]	football

About 'university life'

	Target Cognate			Distractor		
	Russian	Transcription	Translation	Russian	Transcription	Translation
1	student	[stʊˈdʲent]	student	aeroport	[aɪrɐˈport]	airport
2	professor	[praˈfʲesər]	prefessor	brunet	[brʲuˈnʲet]	brunette
3	doktor	[ˈdoktər]	doctor	zebra	[ˈzʲebrə]	zebra
4	universitet	[ʊnʲɪvʲɪrsʲɪˈtʲet]	university	menu	[mʲɪˈnʲu]	menu
5	kyrs	[kurs]	course	pasport	[ˈpaspərt]	passport
6	lekciya	[ˈlʲektsɨjə]	lecture	prezident	[prʲɪzʲɪˈdʲent]	president
7	sistema	[sʲɪˈsʲtʲemə]	system	shokolad	[ʃɨkɐˈlat]	chocolate
8	biblioteka	[bʲɪblʲɪaˈtʲekə]	bibliotheca	djinsy	[ˈdʒɨnsɨ]	jeans
9	semestr	[sʲɪˈmʲestr]	semester	parashut	[pəraˈʂut]	parachute
10	ekzamen	[ɪgˈzamʲɪn]	examination	radiaciya	[rədʲɪˈatsɨjə]	radiation

Appendix B.1 Language History Questionnaire

Date:

Below are questions about your education, profession, and language use. Please answer these questions as completely as possible.

Background:

First name + Surname initial:

Age:

Sex:

What is your level of education (e.g. high school, university degree):

If you are a student. Which degree do you study?

Which language(s) did you study for your A-levels, GCSE?

Were you born in Newcastle?   Yes   No

If yes:

Have you lived in Newcastle since birth? Yes   No

If no:

Where else have you lived?

How old were you when you came to Newcastle?

How long have you been living in Newcastle?

Have you returned to the place of your birth for longer than 6 months (if yes, how long)?
Yes_______   No

Language history:

What is your native language?

Please list any other languages that you know below. For each, rate how well you can use the language on the following scale:

Not Good 1 2 3 4 5 Very Good

Language	Speaking	Listening	Writing	Reading	Grammar	Pronunciation
1						
2						
3						
4						

For the language you listed, please indicate below the place and age at which you learned them, and if applicable, whether you learned them by formal lessons (e.g., at school or a course), or by informal learning (e.g., at home, at work, from friends).

Language	Country	Age	Lessons (yes/no)	Duration of lessons	Informal (yes/no)	Duration of informal learning
1						
2						
3						
4						

For the languages you listed, rate how well you agree with the following statements using the scale:

Language	I like to speak this language	I feel confident using this language	I think it is important to be good at this language
1			
2			
3			
4			

For the languages you listed, which do you use with the following people, for how many hours per day, on what kind of topic and in which place (home, work, etc):

	Language	Hours per day	Topic	Place
Mother				
Father				
Older brother/sister				
Younger brother/sister				
Children				
Other family members				
Housemates				
Partner				
Friends				
Colleagues				

For the languages you listed, which do you use for the following activities and how many hours per day?

Activity	Language	Hours per day
Reading		
Watching TV		
Listening to the radio		
Email, internet		

In general, how well do you *like* to learn new languages?

Dislike 1 2 3 4 5 Like

In general, how *easy* do you find learning new languages?

Difficult 1 2 3 4 5 Easy

Have you ever taken a formal module in Linguistics?

If you have any other remarks about your language history that you think may be important for your ability to use these languages, please feel free to write them here:

Appendix B.2 Advertisement looking for participants

GET £ 10 AMAZON VOUCHER FOR PARTICIPATING IN EXPERIMENT

I’m looking for native speakers of English without knowledge of Slavic languages to take part in a Linguistics experiment. You will be asked to meet with the researcher for 30 minutes over four days. You will receive £10 Amazon voucher on a final day. You will also get an explanation of your learning curve!

On each day you will need to listen to an audio file of an unknown language and then do three listening tasks on an experimenter’s laptop. On the final day you will also be asked to complete a short questionnaire. All testing will take place in a quiet room in Percy building, Newcastle University.

For more information or to arrange your appointments please contact Natalia (n.v.pavlovskaya@ncl.ac.uk)

Appendix B.3 Consent form

CONSENT FORM

School of English Literature,

Language and Linguistics,

Percy Building,

Newcastle upon Tyne,

NE1 7RU, UK

You are asked to participate in this study because you a native speaker of English and your knowledge of Slavic languages fits the criteria for study subjects participating in the researcher's PhD study.

In this study you will be asked to meet with the researcher over four days for a maximum of 30 minutes. On each day you will need to listen over the headphones to an aural input twice and then do three listening tasks on the researcher's laptop. On final day you will also be asked to complete a short questionnaire which is designed to collect bibliographic information relating to your exposure to your second language(s). Your full name will not be recorded or written anywhere; instead a code will be used. You may end participation at any time.

Thank you very much for your participation.

PhD student Natalia Pavlovskaya

AGREEMENT

I agree to participate and allow the recording of my interview and accompanying material to be used for the purpose of this assignment. I understand that I my participation is voluntary and that I have the option of declining to cooperate further at any time during the interview.

Signature of Researcher: ___________________________

Signature of Participant: ___________________________

Dates of Interview: _______________________________

Chapter 9. Appendix B.4 Answers on questionnaire

Participant	Age	Gender	Education	Degree	Which language(s) did you study for your A-levels, GCSE?
Part18	22	female	3d year UG	Speech and Language Sciences	French, Spanish and English
Part12	21	female	3d year UG	English Literature	French, Latin
Part21	19	female	1st year UG	Food & Human Nutrition	Spanish
Part2	19	female	1st year UG	Food & Human Nutrition	German
Part26	18	male	1st year UG	History	Spanish
Part28	19	female	1st year UG	English Literature	French, Latin
Part5	22	female	2d year UG	Spanish & Business	French, Spanish
Part17	20	female	1st year UG	Combined Honours in Linguistics and Spanish	Spanish, French, and Latin
Part22	20	female	2d year UG	Combined Honours in Linguistics and Japanese	German, Spanish, French
Part24	21	female	2d year UG	Combined Honours in Music and Film	Spanish, French, Latin, and Japanese
Part8	20	female	2d year UG	English Literature with Japanese	n/a
Part3	19	female	2d year UG	Linguistics with French	French, German
Part15	32	female	1st year UG	Linguistics with Chinese	none
Part16	19	male	1st year UG	History	French
Part23	18	female	1st year UG	Combined Honours: History, Politics, Business	French
Part11	22	female	3d year UG	Linguistics with Japanese	Japanese
					French
					Creole
Part25	19	male	1st year UG	English Literature and History	French
Part20	21	female	1st year UG	Speech & Language Sciences	Welsh
Part13	18	female	1st year UG	English Language	French
Part7	28	female	MA	Creative Writing	French
Part19	n/a	male	MA	Literature	n/a
Part14	43	male	MA	Creative Writing	
Part27	34	female	MA	Linguistics	Spanish
Part10	23	female	PhD	Linguistics	
Part1	27	male	PhD	Linguistics	French
					German
Part9	37	male	PhD	Drama	French

					German
Part29	49	female	PhD	Linguistics	
Part4	26	female	PhD	English Literature	French

Participant	Born in Newcastle?	Lived in Newcastle since birth?	If no, where else have you lived?
Part18	no	no	Gloucestershire
Part12	no	no	Surrey
Part21	no	no	South Africa; Oxford, UK
Part2	no	no	Lincolnshire
Part26	no	no	Leeds
Part28	no	no	Canterbury
Part5	no	no	Oxfordshire, Somerset
Part17	yes	no	Spain, age 7 to 16
Part22	no	no	Durham
Part24	no	no	San Diego, California & London, UK
Part8	no	no	Turin, Brussels, Abingdon
Part3	no	no	Weybridge birth-2005, Evreham 2005-now
Part15	no	no	Telford, Shropshire
Part16	no	no	Bradford, Leeds
Part23	yes	yes	Newcastle, UK
Part11			
Part25			
Part20	no	no	North Wales
Part13			
Part7	no	no	County Durham, Cambridge, Kingston upon Thames, Hong Kong
Part19	yes	yes	Newcastle, UK
Part14			
Part27	yes	yes	1 year in Spain, 1 year in Japan
Part10			
Part1	no	no	
Part9	no	no	
Part29			Lived and worked in France
Part4	no	no	Wakrfield, Cambridge, Edinburgh

Participant	When came to Newcastle?	How long have you been living in Newcastle?	Returned to the place of your birth for longer than 6 months	Any other languages that you know below?
Part18	19	3.5 years	no	Spanish
				French
Part12	18	3 years	no	French
				Latin
Part21	18	5 months	no	Spanish
Part2	18	6 months	no	
Part26	18	6 months	no	Spanish
Part28	19	5 months	no	French
				Latin
Part5	20	1 year 6 months	no	Spanish
Part17	16	4 years	n/a	Spanish
				French
Part22	n/a	commuting to NCL	n/a	Spanish
				Japanese
				Korean
				German
Part24	18	2 years and 6 months	no	Spanish
				Japanese
Part8	18	2 years	no	French
				Japanese
Part3	18	1 year 6 months	no	French
				German
				Japanese
				Portuguese
Part15	31	5 months	no	Chinese
Part16	19	5 months	no	French
Part23	n/a	n/a	n/a	French
Part11				Japanese
				French
				Creole
Part25				French
				Italian
				Serbo-Croatian
Part20	20	9 months	no	Welsh
Part13				French
Part7	5	23 years	no	French
				Cantonese
Part19				
Part14				French

				Spanish
Part27			yes	Spanish
Part10				n/a
Part1				French
				German
Part9	37	6 months (part-time)	no	French
				German
Part29				French
				Latin
				Russian
Part4	24	2 years	no	German
				French

Participant	Speaking	Listening	Writing	Reading	Grammar	Pronunciation	At what age
Part18	3	3	3	2	2	4	11 y.o.-18y.o.
	2	2	1	2	1	2	
Part12	4	3	3	4	3	4	12y.o.
	1	1	2	3	2	1	16y.o.
Part21	3	3	3	4	3	5	13y.e.-18y.o.
Part2							
Part26	2	2	2	2	2	2	15y.o.
Part28	3	3	3	3	2	3	11y.o.-15y.o.
	2	1	4	5	4	3	11y.o.-15y.o.
Part5	3	4	3	5	3	5	13y.o.-now
Part17	5	5	5	5	5	5	7y.o.-16y.o.
	3	4	3	2	2	2	11y.o.-16y.o.
Part22	4	2	3	4	3	2	11y.o.-now
	3	3	2	4	3	4	18y.o.-now
	1	1	1	1	1	1	20y.o.-now
	1	2	1	2	1	1	16y.o-17y.o.
Part24	4	3	4	5	5	5	10y.o.-18y.o.
	3	3	2	2	3	4	15y.o.-18y.o.
Part8	4	5	4	5	5	5	7y.o.-18y.o.
	3	3	3	3	3	3	18y.o.-now
Part3	4	4	4	5	3	3	7y.o.-now
	3	3	3	4	2	3	11y.o.-18y.o.
	1	1	1	1	1	1	19y.o.-now
	2	2	2	3	2	2	18y.o.-19y.o.
Part15	4	4	2	4	4	4	16y.o.-26y.o.
Part16	2	1	1	2	1	1	8y.o.-16y.o.
Part23	n/a	n/a	n/a	n/a	n/a	n/a	

Part11	4	4	3	3	3	3	
	2	4	1	3	1	4	
	2	3	0	0	2	2	
Part25	4	5	3	4	2	5	
	2	2	1	3	1	3	
	1	1	2	2	1	2	
Part20	5	5	5	5	4	5	from birth
Part13	2	2	1	3	1	2	
Part7	2	2	1	1	1	1	12y.o.-16y.o.
	2	2	1	1	1	1	24y.o.-27y.o.
Part19							
Part14	3	2	2	3	2	3	
	3	2	2	3	1	2	
Part27							
Part10							
Part1	3	3	4	4	4	4	
	2	3	3	3	3	3	
Part9	2	1	1	2	1	2	11y.o.-16y.o.
	1	1	1	1	1	1	12y.o.-14y.o.
Part29	4	4	4	5	4	3	
			2	3	4		
	1	1	1	2	2	1	
Part4	1	1	1	2	2	2	22y.o.-26y.o.
	1	1	1	2	1	2	16y.o.-21y.o.

Participant	Lessons	I like to speak this language	Confident using this language	Important to be good at this language	Like to learn new Languages	Easy learning new languages	Module in Linguistics
Part18	yes	5	3	4	5	3	yes
		4	2	4			
Part12	yes	4	3	3	4	3	
	yes	2	1	2			
Part21	yes	5	3	2	4	3	no
Part2							
Part26	yes	2	3	5	1	1	no
Part28	yes	3	2	5			
	yes	3	2	5	5	1	no
Part5	yes	5	3	5	4	2	no
Part17	no	5	5	5	5	3	yes
	yes	2	2	5			
Part22	yes	5	5	5	5	4	yes
	yes	5	1	5			

	yes	5	1	5			
	yes	1	1	5			
Part24	yes	yes	yes	yes	3	2	no
	yes	yes	no	yes			
Part8	yes	5	4	4	4	2	no
	yes	4	2	3	5	3	yes
Part3	yes	5	4	5			
	yes	4	3	4			
	yes	5	1	5			
	yes	4	2	4			
Part15	no	3	2	2	5	2	yes
Part16	yes	1	1	1	1	1	no
Part23					4	2	
Part11					4	3	yes
Part25	yes				5	4	no
	no						
	no						
Part20	both	5	5	5	3	3	yes
Part13					2	1	yes
Part7	yes	4	2	2	4	2	no
	no	4	2	4			
Part19					5	3	
Part14					4	4	
Part27					5	4	
Part10					2	1	
Part1					3	4	
Part9		4	2	3	4	2	
		2	1	4			
Part29					5	4	
Part4	yes	3	1	5	4	1	
	yes	3	1	5			

Participant	Remarks about language history?
Part18	Studied phonetics so maybe more aware of sound structures of different languages
Part28	Dyslexia & long-term memory loss. Latin helps with English reading/understanding
Part22	Basic Swedish, Finish, Mexican slang
Part24	my father is a professor in Linguistics and can speak Spanish (but I don't speak it with him, just English)
Part16	Reported French, but actually do not speak it with anyone
Part25	I learned foreign language while travelling. I have only been formally tested in French
Part13	Not very good at other languages. Live in a small town where not many languages are heard
Part19	Also, studied customer service at Northumbria Uni. Did a module in 'Language Acquisition and history of English' but can hardly remember anything
Part14	I have a learning difficulty which might affect my ability to learn new languages, but I try to learn a little of many languages
Part1	I learned French and German in School but had no opportunity to use them since then, so my proficiency will have decreased
Part9	I think it helps to be interested in the culture and country
Part29	BA honours in French and Russian, graduated in 1991, haven't used Russian since then. Used French rarely since lived and worked there 1992/93
Part4	Short course in basic German at the Uni of Cambridge (2010/11); short course in German reading (online provided by Durham Uni, 2017)

References

Alderete, J. (1995). ‘Faithfulness to prosodic heads’. Unpublished manuscript, University of Massachusetts.

Alderete, J. (2001). *Morphologically Governed Accent in Optimality Theory.* New York: Routledge.

Allen, M. (2017). 'Between-subjects design'. *The SAGE Encyclopedia of Communication Research Methods*.

Altenberg, E. (2005a). 'The perception of word boundaries'. *Second Language Research*, 21: 325-358.

Altenberg, E. P. (2005b). 'The judgment, perception, and production of consonant clusters in a second language'. *International Review of Applied Linguistics in Language Teaching*, 43: 53-80.

Altenberg, E. P., & Cairns, H. S. (1983). 'The effects of phonotactic constraints on lexical processing in bilingual'. *Journal of Verbal Learning and Verbal Behavior*, 22: 174-188.

Amazon Mechanical Turk. www.mturk.com.

Anderson, J. (1987). ‘The markedness differential hypothesis and syllable structure’. In G. Ioup, & S. Weinberger (Eds.), *Interlanguage Phonology* (pp. 279–291). Rowley, MA: Newbury House.

Archibald, J. (1992). 'Transfer of L1 parameter settings: Some empirical evidence from Polish metrics'. *Canadian Journal of Linguistics*, 37: 301-339.

Archibald, J. (1993). *Language Learnability and L2 Phonology: The Acquisition of Metrical Parameters*. Dordrecht: Kluwer.

Archibald, J. (1994). 'A formal model of learning L2 prosodic phonology'. *Second Language Research*, 10: 215-240.

Arunachalam, S. (2013). 'Experimental methods for linguists'. *Language and Linguistics Compass*, 221–232.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). 'Computation of conditional probability statistics by 8-month-old infants'. *Psychological Science*, 9: 321-324.

Azzopardi, P., & Cowey, A. (1998). 'Visual pattern and motion detection in a hemianopic subject'. *Journal of Cognitive Neuroscience*, 14.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2007). 'Mixed-effects modeling with crossed random effects for subjects and items'. *Journal of Memory and Language*, 59: 390-412.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *Linear Mixed-Effects Models Using Eigen and S4. R Package Version 1.1-8.* Retrieved from http://CRAN.R-project.org/package=lme4

Bernstein, J., & Franco, H. (1996). 'Speech recognition by computer'. In N. J. Lass (Ed.), *Principles of Experimental Phonetics* (pp. 408-434). St. Louis, MO: Mosby.

Bialystok, E. (1979). 'Explicit and implicit judgments of L2 grammaticality'. *Language Learning*, 29: 81-103.

Bialystok, E. (1991). *Language Processing in Bilingual Children.* Cambridge: Cambridge University Press.

Bloom, P. (2000). *How Children Learn the Meanings of Words.* Cambridge, MA: MIT Press.

Bloomfield, L. (1933). *Language.* New York: Henry Holt.

Bond, Z. S. (1999). *Slips of the Ear: Errors in the Perception of Casual Conversation.* San Diego, CA, US: Academic Press.

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2017). 'How many trials does it take to get a significant ERP effect? It depends'. *Psychophysiology*, 55 (6): 13-49.

Brent, M. R. (1999). 'Speech segmentation and word discovery: a computational perspective'. *Trends in Cognitive Sciences*, 3: 294-301.

Brent, M. R., & Cartwright, T. A. (1996). 'Distributional regularity and phonotactic constraints are useful for segmentation'. *Cognition*, 61: 93-125.

Broselow, E., & Finer, D. (1991). 'Parameter setting in second language phonology and syntax'. *Second Language Research*, 7: 35-59.

Brown, R., & Hildum, D. C. (1956). 'Expectancy and the identification of syllables'. *Language*, 32: 411-419.

Carlisle, R. S. (1991). 'The influence of environment on vowel epenthesis in Spanish/English interphonology'. *Applied Linguistics*, 12: 76-95.

Carroll, S. E. (1999). ‘Putting ‘input’ in its proper place’. *Second Language Research*, 15 (4): 337-388.

Carroll, S. E. (2001). *Input and Evidence: The Raw Material of Second Language.* Amsterdam: John Benjamins.

Carroll, S. E. (2004). 'Segmentation: learning how to ‘hear words’ in the L2 speech stream'. *Phillological Society*, 102: 227-254.

Carroll, S. E. (2012). 'Segmentation on first exposure to an L2: evidence for knowledge-driven, topdown processing'. In K. Braunmüller, C. Gabriel, & B. Hänel-Faulhaber (Ed.), *Multilingual Individuals and Multilingual Societies* (pp. 23-45). Amsterdam: Benjamins.

Carroll, S. E. (2013). 'Introduction to the special issue: Aspects of word learning on first exposure to a second language'. *Second Language Research*, 29.

Carroll, S. E. (2014). ‘Processing 'words' in early-stage SLA: a comparison of first exposure and low proficiency learners’. In Z. Han, & R. Rast (Ed.), *First Exposure to a Second Language: Learners' Initial Input Processing* (pp. 107-138). Cambridge: Cambridge University Press.

Carroll, S. E. (1992). 'On cognates'. *Second Language Research*, 8 (2): 93-119.

Carroll, S. E., & Widjaja, E. (2013). 'Learning exponents of number on first exposure to an L2'. *Second Language Research*, (29) 2: 201-229.

Carroll, S. E.; Windsor, J. (2015). ‘Segmental targets versus lexical interference: production of second-language targets on first exposure with minimal training’. In H. Peukert (Ed.), *Transfer Effects in Multilingual Language Development. Hamburg Series in Language Development 4.* Amsterdam: John Benjamins.

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). 'Infants learn phonotactic regularities from brief auditory experience'. *Cognition*, 87: B69-B77.

Charness, G., Gneezy, U., & Henderson, A. (2013). 'Experimental methods: Eliciting risk preferences'. *Journal of Economic Behaviour and Organisation*, 87: 43-51.

Chew, P. (2000). *A Computational Phonology of Russian.* Parkland, FL: Dissertation.Com.

Chomsky, N. (1987). *The Chomsky Reader*. New York: Pantheon Books.

Christiansen, M. H., & Chater, N. (1994). 'Generalisation and connectionist language learning'. *Mind and Language*, 9 (3): 273-287.

Clements, G. N., & Keyser, S. J. (1983). *CV Phonology a Generative Theory of the Syllable.* Cambridge MA: The MIT Press.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Routledge.

Cole, R., & Jakimik, J. (1980). 'A model of speech perception'. In R. Cole (Ed.), *Perception and Production of Fluent Speec* (pp. 133-163). Hillsdale, NJ: Lawrence Erlbaum Associates .

Corder, S. (1967). 'The significance of leamer's errors'. *International Review of Applied Linguistics in Language Teaching*, 5: 161-170.

Cristia, A., Mielke, J., Daland, R., & Peperkamp, S. (2013). 'Similarity in the generalisation of implicitly learned sound patterns'. *Laboratory Phonology*, 4 (2): 259-285.

Crosswhite, K. (2001). *Vowel Reduction in Optimality Theory.* London & New York: Routledge.

Crosswhite, K., Alderete, J., Beasley, V., & Markman, V. (2003). 'Morphological effects on default stress placement in novel Russian words'. In G. G. Tsujimura (Ed.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics* (pp. 151-164). Somerville, MA: Cascadilla Press.

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). 'Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research'. *PLoS ONE*, 8 (3).

Cutler, A. (1990). ‘Exploiting prosodic probabilities in speech segmentation’. In G. T. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* Cambridge: MIT Press.

Cutler, A. (1994). 'Segmentation problems, rhythmic solutions'. *Lingua*, 92: 81-104.

Cutler, A., & Butterfield, S. (1992). 'Rhythmic cues to speech segmentation: evidence from juncture misperception'. *Jornal of Memory and Language*, 31: 218-236.

Cutler, A., & Carter, D. M. (1987). 'The predominance of strong initial syllables in the English vocabulary'. *Computer Speech und Language*, 2: 133-142.

Cutler, A., & Norris, D. (1988). 'The role of strong syllables in segmentation for lexical access’. *Journal of Experimental Psychology*, 14: 113-121.

Davenport, M., & Hannahs, S. J. (2010). *Introducing Phonetics and Phonology* (third edition). London, UK: Hodder Education.

Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell , M. G. (2009). 'Learning and consolidation of novel spoken words'. *Journal of Cognitive Neuroscience*, (4): 803-20.

De Graaff, R. (1997). 'The eXperanto experiment: Effects of explicit instruction on second language acquisition'. *Studies in Second Language Acquisition*, 19 (2): 249-276.

DeCasper, A. J., & Spence, M. J. (1986). 'Prenatal maternal speech influences newborns' perception of speech sounds'. *Infant Behavior and Development*, 9: 133-150.

DeCasper, A., & Fifer, W. P. (1980). 'Of human bonding: newborns prefer their mothers' voices'. *Science*, 208 (4448): 1174–1176.

DeKeyser, R. (1997). 'Beyond explicit rule learning: Automatizing second language morphosyntax'. *Studies in Second Language Acquisition*, 19 (2): 195-221.

DeKeyser, R. (2003). ‘Implicit and explicit learning’. In C. Doughty, & M. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 313-349). Malden, MA: Blackwell.

Docherty, G. J. (1992). *The Timing of Voicing in British English Obstruents.* Berlin: Foris.

Dommergues, J.-Y., & Segui, J. (1989). 'List structure, monotony, and levels of processing'. *Journal of Psycholinguistic Research*, 18(3): 245-253.

Doughty, C. (1991). 'Second language instruction does make a difference: Evidence from an empirical study on SL relativization'. *Studies in Second Language Acquisition*, 13: 431-469.

Duanmu, S. (2000). *The Phonology of Standard Chinese.* Oxford: Oxford University Press.

Echols, C. H. (1993). 'A perceptually-based model of children’s earliest productions'. *Cognition*, 46: 245-296.

Eckman, F. (1977). 'Markedness and the contrastive analysis hypothesis'. *Language Learning*, 27: 315 – 33.

Edelman, S., & Christiansen, M. (2003). 'How seriously should we take Minimalist syntax?' *Trends in Cognitive Sciences*, 7: 60-6.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). 'Speech perception in infants'. *Science*, 171: 303-306.

Eimas, P. D., & Miller, J. L. (1981). 'Organization in the perception of segmental and suprasegmental information by infants'. *Infant Behavior and Development*, 4: 395-399.

Elder, C. (2009). ‘Validating a test of metalinguistic knowledge’. In R. Ellis, S. Loewen, C. Elder, J. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching* (pp. 113-138). Bristol, UK: Multilingual Matters.

Elder, C., & Manwaring, D. (2004). 'The relationship between metalinguistic knowledge and learning outcomes among undergraduate students of Chinese'. *Language Awareness*, 13: 145-162.

Elder, C., Warren, J., Hajek, J., Manwaring, D., & Davies, A. (1999). 'Metalinguistic knowledge: How important is it in studying a language at university?'. *Australian Review of Applied Linguistics*, 22: 81-95.

Ellis, N. (1993). 'Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge'. *European Journal of Cognitive Psychology*, 5: 289-319.

Ellis, N. (2003). ‘Constructions, chunking, and connectionism: The emergence of second language’. In C. A. Doughty, & M. H. Long (Eds.), *Handbook of Second Language Acquisition.* Blackwell Publishing.

Ellis, R. (2004). 'The definition and measurement of explicit knowledge'. *Language Learning*, 54: 227-275.

Ellis, R. (2009). ‘Implicit and explicit learning, knowledge and instruction’. In R. Ellis, S. Loewen, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching.* Bristol: Multilingual Matters.

Ellis, R. (2009). ‘Measuring implicit and explicit knowledge of a second language’. In R. Ellis, S. Loewen, C. Elder, J. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and Explicit*

Knowledge in Second Language Learning, Testing and Teaching (pp. 31-64). Bristol, UK: Multilingual Matters.

Finley, S., & Badecker, W. (2009). 'Artificial language learning and feature-based generalisation'. *Journal of Memory and Language*, 61 (3): 423-437.

Flege, J. (1992). ‘The intelligibility of English vowels spoken by British and Dutch talkers’. In R. D. Kent (Ed.), *Intelligibility in Speech Disorders: Theory, Measurement, and Management* (pp. 157–232). Philadelphia: John Benjamins.

Flege, J. E. (1999). ‘Age of learning and second language speech’. In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis* (pp. 101–131). Mahwah, NJ: Lawrence Erlbaum Associates.

Flege, J. E., Munro, M., & Mackay, I. (1995). 'Factors affecting degree of perceived foreign accent in a second language'. *Journal of the Acoustical Society of America*, 97: 3125–3134.

Folia, V., Udden, J., De Vries, M., Forkstam, C., & Petersson, K. M. (2010). 'Artificial language learning in adults and children'. *Language Learning*, 60: 188-220.

Francis, W. N., & Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar.* Boston, MA: Houghton Mifflin.

Friederici, A. D., & Wessels, M. I. (1993). 'Phonotactic knowledge of word boundaries and its use in infant speech perception'. *Perception & Psychophysics*, 3: 287-295.

Gass, S. M., & Selinker, L. (2008). *Second Language Acquisition: An Introductory Course* (third edition). New York, NY: Routledge.

Gibson, E., & Fedorenko, E. (2010). 'The need for quantitative methods in syntax and semantics research'. *Language and Cognitive Processes*, 28: 1-2, 88-124.

Gluck , M. A., Mercado, E., & Myers, C. E. (2007). *Learning and Memory: From Brain to Behavior.* Worth Publishers.

Goldberg, A. E. (2006). *Constructions at work: the Nature of Generalisation in Language.* New York: Oxford University Press .

Gomez, R. (2002). 'Variability and detection of invariant structure'. *Psychological Science*, 13: 431-436.

Green, D. S., & Sweets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Vol.1). New York: Wiley.

Guion, S. G., Clark, J. J., Harada, T., & Wayland, R. (2003) 'Factors affecting stress placement for English nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words'. *Language & Speech*, 46 (4): 403-27.

Gullberg , M., & Indefrey, P. (2003). *Language Background Questionnaire. Developed in The Dynamics of Multilingual Processing*. (Nijmegen, Max Planck Institute for Psycholinguistics) Retrieved July 07, 2019, from https://www.researchgate.net/...questionnaires.../Lang-Hist-Quest-Engl.pdf

Gullberg, M., Roberts, L., & Dimroth, C. (2012). ‘What word-level knowledge can adult learners acquire after minimal exposure to a new language?’. *International Review of Applied Linguistics*, 50: 239-276.

Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., & Indefrey, P. (2010). 'Adult language learning after minimal exposure to an unknown natural language'. *Language Learning*, 60: 5-24.

Gurevich, N. (2001). 'A critique of markedness-based theories in Phonology'. *Faculty Research and Creative Activity*, 8.

Halle, M. (1959). *The Sound Pattern of Russian: a Linguistic and Acoustical Investigation.* The Hague: Mouton.

Halle, M. (1975). ‘On Russian accentuation’. *Slavonic and East European Journal*, 19: 104-111.

Halle, M. (1997). 'On stress and accent in Indo-European'. *Language*, 73: 215-313.

Halle, M., & Vergnaud, J. R. (1987). *An Essay on Stress Current Studies in Linguistics.* Cambridge, MA: MIT Press.

Han, Z.-H., & Liu, Z. H. (2013). 'Input processing of Chinese by ab initio learners'. *Second Language Research*, 29 (2): 145-164.

Hanulikova, A., McQueen, J. M., & Mitterer, H. (2010). ‘Possible words and fixed stress in the segmentation of Slovak speech’. *Journal of Experimental Psychology*, 63 (3): 555-79.

Hanulikova, A., Mitterer, H., & McQueen, J. (2011). 'Effects of first and second language on segmentation of non-native speech'. *Bilingualism: Language and Cognition*, 14 (4): 506-521.

Harris, J. (1994). *English Sound Structure.* Oxford, UK: John Wiley & Sons Inc

Hart, D. K. (1998). 'Traces of English stress parameters in the Russian of English speakers'. *The Slavic and East European Journal*, 42: 268-282 .

Haselton, M. G., Nettle, D., & Andrews, P. W. (2005). ‘The evolution of cognitive bias’. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology* (pp. 724–746). Hoboken, NJ, US: John Wiley & Sons Inc.

Haspelmath, M. (2006). 'Against markedness'. *Journal of Linguistics*, 42: 25-70.

Haugen, E., & Joos, M. (1972). ‘Tone and intonation in East Norwegian’. In D. Bolinger (Ed.), *Intonation.* Harmondsworth, England: Penguin, Ltd.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). 'The weirdest people in the world?'. *Behavioral and Brain*, 33: 83-135.

Hohne, E. A., & Jusczyk, P. W. (1994). 'Two-month-old infants' sensitivity to allophonic differences'. *Perception & Psychophysics*, 56: 613-623.

Holdnack J. A., & Drozdick L. W. (2010). ‘Auditory memory’. *Using WAIS-IV Clinical Use and Interpretation*.

Hulstijn, J. H. (2005). 'Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction'. *Studies in Second Language Acquisition*, 27 (02): 129-140.

Idsardi, W. (1992). *The Computaion of Prosody.* Cambridge, MA: MIT.

Ioup, G., Boustagui, E., Tigi, M. E., & Moselle, M. (1994). 'Reexamining the critical period hypothesis: a case study of successful adult SLA in a naturalistic environment'. *Studies in Second Language Acquisition*, (16) 1: 73-98.

Ito, K., & Strange, W. (2009). 'Perception of allophonic cues to English word boundaries by Japanese second language learners of English'. *The Journal of the Acoustical Society of America*, 125 (4): 2348-60.

Johnson, J., & Newport, E. (1989). 'Critical period effects in second language learning: the influence of maturational state on the acquisition of ESL'. *Cognitive Psychology*, 21: 60-99.

Johnson, E. K., & Jusczyk, P. W. (2001). 'Word segmentation by 8-month-olds: when speech cues count more than statistics'. *Journal of Memory and Language*, 44: 548-567.

Jones, D., & Ward, D. (1969). *The Phonetics of Russian.* Cambridge: Cambridge University Press.

Jusczyk, P. W. (1985). 'The high amplitude sucking procedure as a methodological

tool in speech perception research'. In G. Gottlieb and N. A. Krasnegor (Eds.), *Measurement of Audition and Vision in the First Year of Postnatal Life: A Methodological Overview* (pp. 195-222). Westport, CT, US: Ablex Publishing.

Jusczyk, P. W., Friederici, A. D., Wessels, J. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). 'Infants′ sensitivity to the sound patterns of native language words'. *Journal of Memory and Language*, 32: 402-420.

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L., Woodward, A., & Piwoz, J. (1992). 'Perception of acoustic correlates of major phrasal units by young infants'. *Cognitive Psychology*, 24, 252-293.

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999a). 'Infants' sensitivity to allophonic cues'. *Perception & Psychophysics*, 61: 1465-1476.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999b). 'The beginnings of word segmentation in English-learning infants'. *Cognitive Psychology*, 39: 159-207.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). 'Infants′ sensitivity to phonotactic patterns in the native language'. *Journal of Memory and Language*, 33: 630-645.

Jusczyk, P., & Aslin, R. (1995). 'Infants' detection of the sound patterns of words in fluent speech'. *Cognitive Psychology*, 1: 1-23.

Jusczyk, P., Cutler, A., & Redanz, N. (1993). 'Infants' preference for the predominant stress'. *Child Development*, 64: 675-687.

Kellerman, E. (1979). 'Transfer and non-transfer: where we are now'. *Studies in Second Language Acquisition*, 2: 37–57.

Kelly, M. H., & Bock, J. K. (1994). 'Stress in time'. *Journal of Experimental Psychology: Human Perception and Performance*, 14: 389-403.

Kirk, N. W., Scott-Brown, K. C., & Kempe, V. (2013). ‘How well can listeners distinguish dialects and unfamiliar languages?'. Paper presented at 54th Annual Meeting of The Psychonomic Society Toronto, Canada.

Klein, W. (1986). *Second Language Acquisition.* Cambridge: Cambridge University Press. .

Krashen, S. (1978). ‘Individual variation in the use of the monitor’. In W. Ritchie (Ed.), *Second Language Acquisition Research.* New York: Academic Press.

Kuhl, P. K. (1985). ‘Methods in the study of infant speech perception’. In G. Gottlieb, & N. A. Krasnegor (Eds.), *Measurement of Audition and Vision in the First Year of Postnatal Life.* Norwood: NJ: Ablex.

Lado, R. (1957). *Linguistics Across Cultures.* Ann Arbor: University of Michigan Press.

Lavitskaya, Y. (2015). ‘Prosodic structure of Russian: A psycholinguistic investigation of the metrical structure of Russian nouns’. Unpublished PhD dissertation, University of Konstanz.

Lebedeva, J. G. (1986). *Zvuki, Udarenie, Intonacija (Sounds, Stress, Intonation).* Moscow: Russkij Jazyk.

Lenneberg, E. (1967). *Biological Foundations of Language.* New York: John Wiley.

Linzen, T., & Gallagher, G. (2017). 'Rapid generalisation in phonotactic learning'. *Journal of the Association for Laboratory Phonology*, 8 (1): 1-32 .

Lust, B. (2006). *Child Language: Acquisition and Growth.* Cambridge University Press.

MacMillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User’s Guide* (second ed.). New York: Psychological Press.

Major, R. C. (2008). ‘Transfer in second langauge phonology’. In J. G. Hansen Edwards, & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 63-95). Amsterdam, The Netherlands: John Benjamins B.V.

Makowski, D. (2018). *Psycho. R Package, Version 0.3.4.* Retrieved from https://CRAN.R-project.org/package=psycho

Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). 'Infants’ recognition of the sound patterns of their own names'. *Psychological Science*, 6: 315-318.

Masny, D. (1987). ‘The role of language and cognition in second language metalinguistic awareness’. In J. P. Lantolf, & A. Labarca (Eds.), *Research in Second Language Learning: Focus on the Classroom* (pp. 59-73). Norwood: NJ: Ablex.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). 'OpenSesame: An open-source, graphical experiment builder for the social sciences'. *Behavior Research Methods*, 44 (2): 314-324.

Mattys, S. J., & Jusczyk , P. W. (2001). 'Phonotactic cues for segmentation of fluent speech by infants'. *Cognition*, 78 (2): 91-121.

Mattys, S. L., Jusczyk, P. W., Luce, A. L., & Morgan, J. L. (1999). 'Phonotactic and prosodic efects on word segmentation in infant'. *Cognitive Psychology*, 38: 465-494.

McQueen, J. M. (1998). 'Segmentation of continuous speech using phonotactics'. *Journal of Memory and Language*, 39: 21-46.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). 'A precursor of language acquisition in young infants'. *Cognition*, 29: 144-178.

Melvold, J. L. (1990). ‘Structure and stress in the phonology of Russian’. Unpublished PhD dissertation, MIT.

Messer, S. (1967). 'Implicit phonology in children'. *Journal of Verbal Learning and Verbal Behavior*, 6: 609-613 .

Miller , J. L., & Liberman, A. M. (1979). 'Some effects of later-occurring information on the perception of stop consonant and semivowel'. *Perception & Psychophysics, 25*: 457-465.

Milroy, T. H. (1909). 'Fatigue studied in reaction time experiments'. *Quarterly Journal of Experimental Physiology*, 2 (3): 277-282.

Mołczanow, J., Domahs, U., Knaus, J., & Wiese, R. (2013). 'The lexical representation of word stress in Russian: evidence from event-related potentials'. *The Mental Lexicon*, 164-194.

Newman, R., Tsay, J., & Jusczyk, P. W. (2003). ‘The development of speech segmentation abilities'. (D. Houston, A. Seidl, G. Hollich, & A. Jusczyk, Editors) Retrieved from Jusczyk Lab Final Report: http://hincapie.psych.purdue.edu/Jusczyk/

Norris, J., & Ortega, L. (2000). 'Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis'. *Language Learning*, 50 (3): 417-528.

Ostapenko, O. (2005). 'The optimal L2 Russian syllable onset'. *LSO Working Papers in Linguistics 5: Proceedings of WIGL (Workshop in General Linguistics)*, 140-151.

Park, E. S. (2011). 'Learner-generated noticing of written L2 input: What do learners notice and why?'. *Language Learning*, 61: 146-186.

Patkowski, M. (1980). 'The sensitive period for the acquisition of syntax in a second language'. *Language Learning*, 30: 449–472.

Perdue, C. (1996). ‘Pre-basic varieties. The first stages of second language acquisition’. *Toegepaste Taalwetenschap in Artikelen*, 55: 135-150.

Pienemann, M. (1999). *Language Processing and Second Language Development: Processability Theory.* Amsterdam: John Benjamins.

Pienemann, M. (2007). ‘Processability theory’. In B. VanPatten, & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 137–154). Mahwah, NJ: Lawrence Erlbaum Associates.

Potter-Collins, A. (2013). *Language in England and Wales: 2011*. Retrieved July 15, 2018, from Office for National Statistics: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language/articles/languageinenglandandwales/2013-03-04

Prasada, S., & Pinker, S. (1993). 'Generalisation of regular and irregular morphological patterns'. *Language and Cognitive Processes*, 8 (1): 1-56.

Proctor, M. I. (2009). ‘Gestural characterization of a phonological class: the liquids’. Unpublished PhD Dissertation, Yale University.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.* Vienna, Austria. Retrieved from http://www.R-project.org/

Rast, R. (1998). *Le Traitement de l’Input dans l’Acquisition du Polonais par des.* Universite´ Paris VIII: Me´moire de DEA.

Rast, R. (1999). 'The first hours of second language acquisition'. *Acquisition et Interaction en Langue E´trange`re 2*, Special Issue: 73-88.

Rast, R. (2008). *Foreign Language Input: Initial Processing.* Clevedon: Multilingual Matters.

Rast, R. (2010). ‘The role of linguistic input in the first hours of adult language learning’. *Language Learning*, 60 (2): 64-84.

Rast, R., & Dommergues, J.-Y. (2003). ‘Towards a characterisation of saliency on first exposure to a second language’. *EUROSLA Yearbook*, 3: 131-156.

Rice, K. (1999). 'Featural markedness in phonology: variation'. *Part 1. GLOT 4.7:3-6. Part 2. GLOT* 4.8:3-7.

Ristin-Kaufmann, N., & Gullberg, M. (2014). 'The effects of first exposure to an unknown language at different ages'. *Bulletin VALS-ASLA*, 99: 17-29.

Robinson, P. (1996). 'Learning simple and complex rules under implicit, incidental rule-search conditions, and instructed conditions'. *Studies in Second Language Acquisition*, 18: 27-67.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). 'Statistical learning by 8-month-old infants'. *Science*, 274 (5294): 1926-8.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). 'Word segmentation: the role of distributional cues'. *Journal of Memory and Language*, 35: 606-621.

Sanders , L. D., & Neville, H. J. (2000). 'Lexical, syntactic, and stress-pattern cues for speech segmentation'. *Journal of Speech, Language, and Hearing Research*, 43: 1301–1321.

Sanders, L. D., & Neville, H. J. (2002). 'Speech segmentation by native and non-native speakers: the use of lexical, syntactic, and stress-pattern cues'. *Journal of Speech, Language, and Hearing Research*, 45 (3): 519-530.

Schwartz, B., & Eubank, L. (1996). 'What is the ‘L2 initial state’?'. *Second Language Research*, 12 (1): 1-5.

Selinker, L. (1972). 'Interlanguage'. *International Review of Applied Linguistics*, 10: 209–231.

Selkirk, E. (1984). ‘On the major class features and syllable theory’. In M. Aronoff, & R. T. Oehfle (Eds.), *Language Sound Structure* (pp. 107-136). Cambridge: MA: MIT Press.

Sereno, J. A. (1986). 'Stress pattern differentiation of form class in English'. *Journal of the Acoustical Society of America*, 79: 36.

Sharoff, S. (2002). 'Meaning as use: exploitation of aligned corpora for the contrastive study of lexical semantics'. *Language Resources and Evaluation Conference (lrec02)*, (pp. 447–452).

Sheldon, A., & Strange, W. (1982). 'The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception'. *Applied Psycholinguistics*, 3: 243–261.

Shoemaker, E., & Rast, R. (2013). ‘Extracting words from the speech stream at first exposure’. *Second Language Research*, 29 (2): 165-183.

Shook, D. J. (1994). 'FL/L2 reading, grammatical information, and the input-to-intake'. *Applied Language Learning*, 5: 57-93.

Singmann, H., Bolker, B., & Westfall, J. (2015). *Afex: Analysis of Factorial Experiments. R Package, Version 0.14-2.* Retrieved from https://CRAN.R-project.org/package=afex

Skirgård, H., Roberts, S., & Yencken , L. (2017). 'Why are some languages confused for others? Investigating data from the Great Language Game'. *PLoS ONE*, 12 (4).

Smith, N. V. (1973). *The Acquisition of Phonology.* London: Cambridge University Press.

Snow, C., & Hoefnagle-Hohle, M. (1978). 'The critical age for second language acquisition: evidence from second language learning'. *Child Development*, 49: 1114-1128.

Stanislaw, H., & Todorov, N. (1999). 'Calculation of signal detection theory measures'. *Behavior Research Methods, Instruments, and Computers*, 31 (1): 137-49.

Swingley, D., & Aslin , R. N. (2000). 'Spoken word recognition and lexical representation in very young children'. *Cognition*, 76 (2): 147-66.

Timberlake, A. (2004). *A Reference Grammar of Russian.* Cambridge: Cambridge University Press.

Tomasello, M. (2000a). *The Cultural Origins of Human Cognition.* Cambridge, MA: Harvard University Press.

Tomasello, M. (2000b). “Do Young Children Have Adult Syntactic Competence?”. *Cognition* 74: 209-253.

Trapman, M. (2007). ‘Phonotactic constraints on consonant clusters in second language acquisition’. Unpublished Master’s thesis, Utrecht University.

Tsay, J., & Jusczyk, P. W. (2003). 'Detection of words in fluent Chinese by English-acquiring and Chinese-acquiring Infants'. (D. Houston, A. Seidl, G. Hollich, E. Johnson, & A. Jusczyk, Editors) Retrieved from Jusczyk Lab Final Report: http://hincapie.psych.purdue.edu/Jusczyk.

Turk, A., Jusczyk, P. W., & Gerken, L. (1995). 'Do Englisli-learning infants use syllable weight to determine stress?'. *Language and Speech*, 38: 143 - 158.

Tversky, A., & Kahneman, D. (1974). 'Judgement under uncertainty: Heuristics and biases'. *Science*, 185 (4157): 1124–1131.

Vainikka, A., & Young-Scholten, M. (1994). ‘Direct access to X’-theory: Evidence from Korean and Turkish adults learning German’. In T. Hoekstra, & B. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar* (pp. 265-316). Amsterdam: John Benjamins.

Vainikka, A., & Young-Scholten, M. (1994). 'The early stages in adult L2 syntax: additional evidence from Romance speakers'. *Second Language Research*. 12: 140–176.

Vainikka, A., & Young-Scholten, M. (1998). ‘The initial state in the L2 acquisition of phrase structure’. In S. Flynn, G. Martohardjono, & W. O’Neil (Eds.), *The Generative Study of Second Language Acquisition* (pp. 17-34). Mahwah: NJ: Lawrence Erlbaum Associates.

van de Weijer, J. (1998). ‘Language input for word discovery’. In *MPI Series in Psycholinguistics (No. 9).* Nijmegen, The Netherlands.

VanPatten, B. (2000). ‘Thirty years of input’. In M. Morris, C. Anderson, C. Klee, & E. Tarone (Eds.), *Social and Cognitive Factors in Second Language Acquisition: Selected Proceedings of the 1999 Second Language Research* (pp. 287-311). Somerville: Cascadilla Press.

Vitevitch, M. S., & Luce, P. A. (1999). 'Probabilistic phonotactics and neighborhood activation in spoken word recognition'. *Journal of Memory and Language*, 40: 374-408.

Vitevitch, M. S., & Luce, P. A. (2004). 'A Web-based interface to calculate phonotactic probability for words and nonwords in English'. *Behavior research methods, instruments, & computers: a journal of the Psychonomic Society*, 36 (3): 481-7.

Vitevitch, M. S., Luce , P. A., Charles-Luce, J., & Kemmerer, D. (1997). 'Phonotactics and syllable stress: implications for the processing of spoken nonsense words'. *Language Speech*, 47-62.

Weber, A. (2000). 'The role of phonotactics in the segmentation of native and non-native continuous speech'. In A. Cutler, M. McQueen, & R. Zondervan (Ed.), *Proceedings of SWAP, Workshop on Spoken Word Access Processes* (pp. 143-146). Nijmegen: MPI for Psycholinguistics.

Weber, A., & Cutler, A. (2006). 'First-language phonotactics in second-language listening'. *The Journal of the Acoustical Society of America*, 1: 597-607.

Werker, J. F., & Tees, R. C. (1983). 'Developmental changes across childhood in the perception of non-native speech sounds'. *Canadian Journal of Psychology*, 37 (2): 178–286.

Woodward, J. Z., & Aslin, R. N. (1990). ‘Segmentation cues in maternal speech to infants’. Paper presented at the meeting of the International Conference on Infant*.* Montreal, Quebec.

Yang, L., & Givon, T. (1997). 'Benefits and drawbacks of controlled laboratory studies of second language acquisition: The Keck second language learning'. *Studies in Second Language Acquisition*, 19 (2): 173-193.

Zaliznjak, A. (1977). *Grammatičeskij Slovar’ Russkogo Jazyka (Russian Grammatical Dictionary).* Moscow: Russkij Jazyk.

Zaliznjak, A. (1985). *Ot Praslavjanskoj Akcentuacii k Russkoj (From Old Russian to Modern Russian Accentuation).* Moscow: Nauka.