UNIVERSITY OF BERGEN

The Effect of Distribution of Exposure: An Experimental Study
Within the Statistical Learning Paradigm

Kristina Osebakken & Iselin Partee

Master’s Programme in Health Sciences – Logopedics
Faculty of Psychology,
Department of Biological and Medical Psychology
Spring 2017
Acknowledgements

The present thesis is made as a completion of the speech and language therapy masters’ degree at the University of Bergen (UiB), Norway. The experiment was conducted at UiB, as a part of a larger research project lead by Dr. Arve Asbjørnsen. A respective number of people have given us academically and practically support. We would like to thank our supervisor Dr. Arve Asbjørnsen at the University of Bergen and assistant supervisor Sunniva Sørhus Eidsvåg. Thank you for the engagement and fast and constructive feedback throughout this process. We also want to thank all participants that took part in the pilot study and the experiment. Without your contribution the experiment would not have been possible. Furthermore, we want to thank UiB for the financial support given to Sunniva Sørhus Eidsvåg, also making the completion of the study possible.

Bergen, May 2017

Kristina Osebakken & Iselin Partee
Table of Contents

Abstract ........................................................................................................................................... 5
Sammendrag ..................................................................................................................................... 6
Theoretical Background .................................................................................................................. 7
  Learning and Language Acquisition ............................................................................................. 7
  Implicit Learning and Statistical Learning .................................................................................... 9
    Implicit Learning ......................................................................................................................... 10
    Statistical Learning ................................................................................................................... 10
  Natural Language Compared to Artificial Language ................................................................. 13
Distribution of Exposure During Learning .................................................................................. 15
  Awareness and Statistical Learning .............................................................................................. 18
  Individual Differences ................................................................................................................ 20
  Clinical Implications of Statistical Learning .............................................................................. 23
Research Questions and Hypotheses ............................................................................................. 25
Methodology ................................................................................................................................... 26
  Research Design ........................................................................................................................ 26
  Participants .................................................................................................................................. 27
  Material and Procedure ............................................................................................................. 29
    Pilot Study ................................................................................................................................. 30
    Recruitment ............................................................................................................................ 30
    Material .................................................................................................................................... 31
      Familiarization Phase .............................................................................................................. 33
      Test Phase .............................................................................................................................. 35
      Self-rating Questionnaire ....................................................................................................... 36
    Procedure ................................................................................................................................. 37
DISTRIBUTION OF EXPOSURE AND SUBCATEGORY LEARNING

Data Processing and Analyses……………………………………………………………38
Analysis of Variance……………………………………………………………………41
Post Hoc Tests……………………………………………………………………………42
Correlation…………………………………………………………………………………42
Ethical Considerations……………………………………………………………………43
Concluding Remarks……………………………………………………………………45
References…………………………………………………………………………………47

Article: The Effect of Distribution of Exposure on Unguided Subcategory Learning in Adults

Appendix A: Self-rating Questionnaire
Appendix B: Consent Form
Appendix C: Instructions on the Computer Screen
Appendix D: Information Form
Abstract

The present study was designed to investigate whether distribution of exposure might facilitate the learning of noun marking of gender subcategories of an unfamiliar, natural language (Russian). Forty graduate students (20 males and 20 females) participated in a between groups experiment where they were familiarized and tested for learning gender marking of Russian nouns. The stimuli was delivered either as three cycles consecutively of 64 items (in total 192 items) before the test phase (massed condition) or divided into three cycles of 64 items separated by a test phase (distributed condition). Learning was operationalized as significantly higher acceptance rate of grammatical items compared to ungrammatical items. The findings revealed that the two conditions produced the learning as measured by higher acceptance of grammatical items than ungrammatical items. The two conditions did not differ in learning outcome. However, the participants’ response time and acceptance rate could indicate that the two conditions triggered different learning strategies. In addition, the measurement of awareness suggested that the two conditions gave similar awareness during the experiment while also reporting recognition of grammatical items more than ungrammatical items.

Keywords: language acquisition, statistical learning, natural language, distribution of exposure
Sammendrag

Denne masteroppgaven ble utviklet for å undersøke om distribuering av eksponering kan lette tilegnelse av kjønnsunderkategorier i et ukjent, naturlig språk (russisk). Førti studenter (20 menn og 20 kvinner) deltok i en mellomgruppedesign hvor de ble kjent og testet for læring av kjønnsmarkering av russiske substantiver. Stimulusmaterialet ble levert enten som en blokk på 192 elementer før testfasen (masse tilstand) eller delt inn i tre blokker med 64 elementer separert av en testfase (distribuert tilstand). Læring ble operasjonalisert som graden av aksept av de grammatiske elementene og avvisning av ugrammatiske elementer under test fasen. Resultatene viste at de to gruppene gav læringen som målt ved aksept av grammatiske elementer og avvisning av ugrammatiske elementer. De to gruppene var ikke forskjellige i læringsutbytte. Deltakernes responstid og aksepteringsrate kan imidlertid indikere at de to betingelsene utløste forskjellige læringsstrategier. I tillegg forstod måling av bevissthet at gruppene hadde liknende bevissthet under forsøket, mens de på samme tid viste en tilbøyelighet til å godta grammatiske elementer mer enn ugrammatiske elementer.

Nøkkelord. Språktilegnelse, statistisk læring, naturlig språk, distribuering av eksponering
Theoretical Background

The present study was designed to investigate whether distribution of exposure might facilitate learning the marking of noun gender subcategories in an unfamiliar, natural language (Russian). Our main objective was to find out if the participants would have an effect of learning when it comes to distribution of exposure. Further, we wanted to see if the distribution of exposure would make a difference in learning the marking of noun gender subcategories. In addition, we thought it would be interesting to investigate if the participants reported having any awareness or engaged in any explicit strategies.

Following, we will present relevant theoretical background, methodological aspects, fundamental research issues and ethical considerations in relation to the experiment.

Learning and Language Acquisition

The process in which children learn a language and accomplish with ease is unknown, a range of language and cognitive processes needs to be integrated in an effortless way (Baird, 2008; Kuhl, 2010). The methods infants use when learning a language are complex and multi-modal (Kuhl, 2010). An infant can discover sounds and words used in their particular language(s) as well as understand the words that induce meaning and use these words to convey their thoughts and desires (Kuhl, 2010). Humans’ capacity for language acquisition is a complex process and the underlying mechanisms and exact nature has led to a debate on nature versus nurture by strong proponents of nativism (Chomsky, 1959) and learning (Skinner, 1959). The debate is whether the language acquisition process could be explained by general learning mechanisms or that learning is relegated to the sidelines. There is extensive empirical evidence that humans are equipped with domain-general learning mechanisms (Saffran, 2003), however it is undeniable that experience-dependent mechanisms are also required for the acquisition of language (Saffran, Aslin, & Newport, 1996). Considerable research has demonstrated that infants can track statistical properties of
language, which is an argument in favor of learning-based language acquisition (Gómez & Lakusta, 2004; Saffran, 2003; Saffran, Aslin, et al., 1996). General learning mechanisms engage on all forms of information in the environment, including input of linguistic nature. Most of the research on such learning mechanisms is gathered by the implicit learning tradition (Reber, 1967) and more recent statistical learning literature (Saffran, Aslin, et al., 1996). Within these traditions the learning process that results from unguided exposure to structured input is examined. To a certain degree, research in these traditions are able to simulate the complex process infants and adults encounter when acquiring a first language or an unfamiliar second language (Gómez & Gerken, 2000; Robinson, 2010).

A fundamental assumption that the ability to learn language is inherent and unique from other cognitive abilities is common in the nativist perspective (Gómez & Gerken, 2000). The nativists further postulate that humans are born with an innate knowledge about structure and grammar and do not require learning. To learn something, one must select the correct structure from an infinite number of potential structures present in any set of data, without the help of feedback (Chomsky, 1965). This might explain how infants are able to acquire a language and rapidly discover the underlying structure of this extensive system with no guidance to clarify the fundamental principles of a language (Chomsky, 1965). Previous studies have shown that the average person can, through auditory input, recognize elements that signal word classes (e.g., verbs, nouns, articles) and identify the underlying “rules” as concerns how these word classes can be paired (Frigo & McDonald, 1998; Gerken, Wilson, & Lewis, 2005; Gómez & Lakusta, 2004; Richardson, Harris, Plante, & Gerken, 2006).

Nevertheless, learning-oriented theories challenged the nativist view by suggesting that a powerful innate learning mechanism to easily obtain knowledge explains acquisition of language (Saffran, 2003). This mechanism needs no explicit guidance regarding any systems or other regularities in the environment (Saffran, 2003). The constrained statistical learning
Framework suggests that learning is central to language acquisition and that the specific nature of language learning explains similarities across different languages (Saffran, 2003). This theory suggests that statistical learning operates on regularities and facilitates processes that vary (i.e., word segmentation, vocabulary learning and syntax) (Finn & Hudson Kam, 2008; Rowland & Pine, 2000; Yu, 2008). Those who hold a different opinion to the nativist perspective claim humans have powerful general cognitive and perceptual mechanisms for learning, which are not restricted to the acquisition of language (Gómez, 2006; Saffran, 2003).

**Implicit Learning and Statistical Learning**

Since most of the research on input of linguistic nature and learning mechanisms are gathered by the implicit learning tradition and the statistical learning tradition, the present experiment was designed within the framework of these paradigms. Research concerning both implicit and statistical learning has been developed into major paradigms in cognitive psychology and developmental psychology respectively (for an overview, see Gómez, 2007; Perruchet, 2008). The two approaches concentrate on how we acquire information from the environment and both rely mostly on the use of artificial grammar (Hamrick & Rebuschat, 2011). The experiments typically have subjects exposed to stimuli generated by an artificial system and then test to determine what they have learned. Based on these similarities, Perruchet and Pacton (2006) suggested that implicit learning and statistical learning exemplify two approaches to one phenomenon. This can be a potential theoretical challenge for future studies as the two approaches can lead to different interpretations of data. Other researchers go as far has combining the two approaches in one name: implicit statistical learning (e.g. Conway & Christiansen, 2006; Emberson, Conway, & Christiansen, 2011; Kidd, 2012; Yim & Rudoy, 2013). Since implicit learning and statistical learning are so intertwined, a brief overview will be presented of both traditions.
**Implicit learning.** In a broad sense implicit learning can be said to be the ability to derive complex information without the consciousness of what has been learned (Cleeremans, Destrebecqz, & Boyer, 1998). The research tradition emerged with the seminal study by Reber (1967) on artificial grammar learning and has since then been investigated by different paradigms, including the artificial grammar learning, sequence learning and dynamic system control (for review, see Cleeremans, et al., 1998). Despite the progress within recent research, the field still suffers from a number of unresolved empirical and theoretical issues (Frensch & Runger, 2003). Firstly, because of arguments due to differences in the understanding of the term, there is no single existing definition. Secondly, results concerning the role of attention in implicit learning are contradictory. Finally, the exact relation between learning and awareness is very much unknown (Frensch & Runger, 2003). The present experiment followed the general procedure of the artificial grammar learning paradigm since it was the most relevant and similar. Most experiments within this paradigm include a familiarization phase where participants are exposed to sequential stimuli that follow a fixed pattern. The participants are then tested with stimuli presented that either share the regularities of the familiarization stimuli or in some way violates the underlying structure (Cleeremans, et al., 1998). The main differences from the artificial grammar learning paradigm is that in the present study we use natural language auditory stimuli as opposed to artificial language material.

**Statistical learning.** The field of language acquisition has been greatly influenced by discoveries of infant and adult learners ability to gather information from speech or constant input and using this information to estimate distinct structures (Bulgarelli & Weiss, 2016; Saffran, Aslin, et al., 1996). The increasing interest in empiricist approaches to language acquisition has made statistical learning, i.e., the ability to identify and take advantage of statistical structure in the environment, develop in a positive direction (Rebuschat &
DISTRIBUTION OF EXPOSURE AND SUBCATEGORY LEARNING

According to Erickson and Thiessen (2015), statistical learning refers to learning on the basis of some conditions of the statistical structure of elements of input in our environment, primarily their frequency, variability, distribution, and co-occurrence probability. The detection, tracking and extraction of these structural regularities appear to be done without awareness, happen during exposure to stimuli and takes place without instruction (Aslin & Newport, 2012; Erickson & Thiessen, 2015).

The early research on statistical learning has been important for demonstrating that infants are able to make extensive use of environmental cues when acquiring language (Ellis, 2006; Rebuschat & Williams, 2011). In their seminal study, Saffran, Aslin, et al. (1996), exposed 8-month-old infants to a continuous stream of speech syllables from an artificial language to see whether they could use statistical information to solve the problem of word segmentation, i.e., discovering word boundaries in continuous speech. The researchers found that the infants differentiated between familiar and unfamiliar stimuli, suggesting that infants are highly sensitive to statistical information and can use this information to accomplish complex learning tasks (Saffran, Aslin, et al., 1996). Subsequent research has shown that the statistical learning ability is sustained throughout adulthood (e.g. Saffran, Newport, & Aslin, 1996) and that statistical learning is not limited to the task of word segmentation (Saffran, Aslin, et al., 1996).

Although statistical learning and rule learning has been claimed to be two separate mechanisms, Aslin and Newport (2012) reviewed evidence and brought together a perspective that argues in favor of one single mechanism of statistical learning that accounts for both the generalization to novel instances and learning of the input stimuli. Neuro-scientific evidence has suggested that statistical learning begins as soon as an individual is exposed to stimuli that contain probabilistic associations, and that this continues in the absence of conscious awareness (Arciuli & Simpson, 2011; Turk-Brown, Scholl, Chun, & Johnson, 2009).
The connection between language universals and statistical learning is under debate. As mentioned, statistical learning is described as a very powerful and domain-general mechanism that is available early in development to infants who are not instructed about how to work out complex learning tasks (Aslin & Newport, 2012; Saffran, 2003; Saffran, Aslin, et al., 1996). Statistical learning research has identified a vigorous mechanism that disposes distributional information across development as well as in different human domains (Culbertson, Smolensky, & Legendre, 2011; Gerken, 2006; Reeder, Newport, & Aslin, 2013). However, further research is required to fully understand the principles underlying such learning results. Statistical learning is something more than just an accurate reproduction of the received input. The process of learning acquisition is influenced by many memory and perceptual forces which not only applies to languages, but also to patterns that are nonlinguistic (Aslin & Newport, 2012).

Some substantial differences are reported between implicit and statistical learning research despite there being an overlap (Hamrick & Rebuschat, 2011). For instance, the most distinctive feature of statistical learning research is the manipulation of statistical information in the input (Hamrick & Rebuschat, 2011). In implicit learning studies this facet is typically not present. Additionally, in statistical learning research the focus is commonly on how to acquire linguistic information, while implicit learning research concentrates on information in general (Hamrick & Rebuschat, 2011). However, more recently the statistical learning paradigm has moved from only investigating early language acquisition to investigate other domains, borrowing concepts from implicit learning tradition for instance by using finite-state grammars to generate stimuli (Perruchet & Pacton, 2006). Lastly, it is widely acknowledged that statistical learning can occur by chance, i.e. without conscious intention to learn, subjects can acquire the statistical structure of language, making this statistical process analogous to that of implicit learning (Hamrick & Rebuschat, 2011). One possible description of the two
paradigms could be that ‘implicit learning’ is a more general term, while ‘statistical learning’ is a form of learning that happens implicitly. With this description in mind, the present experiment is placed within the statistical learning framework in a broad sense in that we examine the learning that results from unguided exposure to structured input.

**Natural Language Compared to Artificial Language**

How the input with statistical information is carefully manipulated might be one of the most distinctive features when it comes to the method used to investigate in statistical learning research (for a review, see Gómez & Gerken, 2000). The input created for these studies typically establish the defining feature of an artificial language (e.g., Gómez, 2002; Gómez & Lakusta, 2004; Mintz, 2002; Saffran, Aslin, et al., 1996). Artificial language has clear advantages because it prevents the learners from benefiting from prior experience with closely related languages or the language itself (Erickson & Thiessen, 2015) as well as allowing a systematic manipulation of the different variables that play into language acquisition (Gómez & Gerken, 1999). While artificial languages provide the researcher with a high degree of control, they also provide simplified linguistic input. As a result, artificial languages lack the natural variability and potential cues, such as pitch, stress and rhythm, which are present in rich and complex natural languages (Aslin & Newport, 2012; Erickson & Thiessen, 2015). This may cause the learning task to differ in important ways from learning in natural languages (Hay, Pelucchi, Estes, & Saffran, 2011).

The importance of reducing the differences between statistical learning and real language acquisition has been emphasized by Romberg and Saffran (2010). Interestingly, Erickson and Thiessen (2015) have explored the feasibility of statistical approaches for natural language acquisition that could make it possible to close this gap. They indicate that the most critical difference between artificial and natural language stimuli used in studies concerning statistical learning, are the different amounts of acoustic variation that exist in a
real language. These acoustic variations can occur in many forms, similar to each phoneme in natural language somewhat varies as a function of the phonemes that predicts and follows it (Erickson & Thiessen, 2015). This kind of variability is not always included in artificial language stimuli. Artificial languages have less complexity, distribution and frequency of words, making the experimenter able to limit on the input the participants are exposed to. This way, learning can be attributed more solely to the cues directly under experimental control. The same does not apply to natural input, as it does not contain the same number of repetitions as artificial language (Erickson & Thiessen, 2015). Still, a noun gender marking system that is present in many natural languages differ conceptually from the artificial grammars used by previous studies (Gómez, 2002; Torkildsen, Dailey, Aguilar, Gómez, & Plante, 2013). Erickson and Thiessen (2015) suggest that future studies on statistical learning should address the ecological validity of the experiment. The downside of natural language stimuli is that it presents multiple and occasionally unnecessary cues to language structure (Braine, 1992). However, mechanisms that could be used in natural language learning have also been proven to be active in adults in learning tasks using artificial language (Mintz, 2002).

To increase the ecological validity of the present experiment, natural language stimuli recorded by native Russian speakers was used instead of an artificial language. The stimuli material consisted of a subset of the Russian gender marking system: feminine and masculine nouns marked for gender by double suffixes. It was organized into subcategories since the input of typical artificial language often sounds monotone and synthesized (Corbett, 1991; Romberg & Saffran, 2010). The gender markings were the stem ending and the case inflection (e.g., Rastochi - tel+ vem, underlined letters being the case inflection). The feminine suffixes used were -k+oj and -k+u, and the masculine suffixes were -tel+yem and -tel+ya. The participants had to detect the relationship between pairs of interchangeable suffixes (i.e. –
koj/ku, -te/yta/-te/yem) to learn the subcategories of masculine and feminine suffixes. Category learning studies have found that presence of multiple cues (e.g., -t el + ya and -k+u) can be helpful in the process of learning (Frigo & McDonald, 1998; McDonald & Plauché, 1995; Mintz, 2002). The effect of excessive cues has also been shown in previous studies with similar Russian stimuli as well as cross-item associations being harder to learn than adjacent dependencies (Eidsvåg, Austad, Plante, & Asbjørnsen, 2015; Gerken, et al., 2005; Richardson, et al., 2006). The adjacent contingency represented by double-marked words provides a more straightforward test of whether learning can withstand the known and damaging effect of inconsistent input on learning. In accordance with previous studies, the double-marked words are considered easier to learn than single-marked (Eidsvåg, et al., 2015), leading to the omission of the single-marked words in the present experiment.

**Distribution of Exposure During Learning**

Research on treatment schedule and language acquisition shows that distribution of exposure in treatment situations can be a potential contributor to outcomes of language learning (Meyers-Denman & Plante, 2016). In the last several years, schedule and intensity of treatment has been an increasing topic within this research field. The distribution of exposure stimuli has been identified as a potential influence for the outcomes on treatment efficiency and language acquisition (Meyers-Denman & Plante, 2016). Within the field of language learning, the distribution of exposure is usually distinguished by the deliverance of doses (Meyers-Denman & Plante, 2016). The stimuli can either be delivered clustered within a compressed period of time (massed), or delivered over a longer period of time (distributed). The number of times the treatment is delivered within a timeframe is known as dose frequency, and the period of time the whole treatment has been delivered can be referred to as total duration of intervention (Warren, Fey, & Yoder, 2007). Both experimental and treatment studies demonstrate contrasting results in relation to distribution of exposure. It can therefore
be possible that either massed or distributed input delivery will have a greater impact on learning abilities (Meyers-Denman & Plante, 2016). The present experiment aims to determine whether a beneficial effect of distribution of exposure could be extended to the learning of noun gender subcategories in an unfamiliar, natural language. A small number of studies have investigated this theory in a context of word learning using different distributions of exposure (distributed and massed) (Ambridge, Theakston, Lieven, & Tomasello, 2006; Janiszewski, Noel, & Sawyer, 2003; Meyers-Denman & Plante, 2016; Proctor-Williams & Fey, 2007; Riches, Tomasello, & Conti Ramsden, 2005; Ukrainetz, Ross, & Harm, 2009; Vlach, Sandhofer, & Kornell, 2008).

An advantage of distributed learning is that it becomes more challenging for memory retrieval (Schmidt & Bjork, 1992). When the items to be learned are more separated in time, the change of context can influence the learning by retrieval of cues or support the encoding with additional cues (Glenberg, 1979; Meyers-Denman & Plante, 2016). With distributed learning, variability of the coding occurs naturally in the trials and variability will also be more likely to occur in the context of learning. Learning delivered over a period can give greater variability of the contextual cues during the process of learning. Experimental studies have shown that distributed and more frequent doses of treatment promote language learning outcomes for young children, compared to longer and sparser treatment doses, which may indicate that a distributed exposure provides more confident outcomes in treatment situations than massed (e.g. Ambridge, et al., 2006; Riches, et al., 2005; Vlach, et al., 2008). However, the advantage of a distributed learning schedule is highly sensitive to both stimuli content and the presentation context (Janiszewski, et al., 2003). Another concern in relation to distributed learning is that the participants can be more likely to forget the previous stimuli before they are presented with the next ones. However, Vlach, et al. (2008) demonstrated that the benefit of distributed learning trials occurred even with the task being more difficult. A
longer gap between the presentation of the items can still be helpful for the memory since it can be more difficult and therefore require more practice (e.g. Vlach, et al., 2008).

Controversy, Barratt, Littlejohns, and Thompson (1992) found that an intensive massed schedule led to greater improvements in expressive language outcomes compared to a distributed learning schedule. 2- to 5-year old children had an overall treatment duration controlled within approximately 6 months. The treatment was admitted either four times a week in a 3-week period within each of two 3-month cycles or one time weekly for 6 months. The authors concluded that the massed treatment led to greater benefits than the distributed treatment, even though the study lacked strong experimental control over the dose frequency and related spacing of treatment sessions (Barratt, et al., 1992).

Other studies have had results that indicate no difference between massed and distributed group performance (Meyers-Denman & Plante, 2016; Proctor-Williams & Fey, 2007; Ukrainetz, et al., 2009). In the study by Meyers-Denman and Plante (2016) 16 children with specific language impairment (SLI) participated in a five week intervention to target grammatical morphology. The children got either one massed treatment session of 30-minutes per day or they had a distributed treatment of three 10-minute sessions within a four-hour period per day. Their progress was assessed three times weekly. The findings indicated that the effect of treatment was significant regardless of distribution condition (Meyers-Denman & Plante, 2016). Congruent with Meyers-Denman and Plante (2016), Proctor-Williams and Fey (2007) had a similar conclusion when it came to children with SLI learning recast densities of novel past tense verbs. Moreover, Ukrainetz, et al. (2009) examined the effects of two intervention schedules on at-risk kindergartners by training phonemic awareness. The findings from this study indicated no difference in performance between the two groups, except for one subskill that favored the individuals that were trained over a longer period of time (Ukrainetz, et al., 2009). Still the main conclusion was that distribution of exposure did
not have any big impact on the results, which shows that results from continuous weekly treatment were similar to those from short and intense treatment (Ukrainetz, et al., 2009).

The abovementioned studies show that there seems to be some disagreement on what kind of distribution of exposure that is most beneficial for learning. Both clinical and experimental studies show disparate results when it comes to distribution of exposure and learning. One can see how either one of massed or distributed treatment can lead to exceptional learning effect, depending on the goal of the treatment. Clinicians may have some flexibility in terms of the dose schedule they use to deliver this treatment in an evidence-based manner (Meyers-Denman & Plante, 2016).

**Awareness and Statistical Learning**

In order for us to be conscious of knowledge, our mental state needs to affirm that the knowledge is present (Dienes, 2011). However, as previously mentioned, the exact relationship between learning and awareness is very much unknown (Frensch & Runger, 2003). The way in which we come to learn about the structure of complex environments is intimately linked to the conscious-unconscious distinction (Dienes, 2011). While implicit learning was previously considered a passive, automatic and unconscious process (c.f., Reber, 1967), it is now widely recognized that a fundamental aspect of human cognition is the process of acquiring unconscious knowledge (Rebuschat, Hamrick, Riestenberg, Sachs, & Ziegler, 2015). There is no longer any doubts whether consciousness is involved but rather the degree to which it is involved (Cleeremans, et al., 1998; Frensch & Runger, 2003).

Williams (2005) presented support for the assumption of learn without awareness, in contrast to an earlier study by Leow (2000), who reported that awareness did not appear to play an important role in second or foreign language development. The difference between the two studies can be explained by methodological differences where the research designs measured unawareness at different stages of the learning process. Rebuschat, et al. (2015)’s
experiment sought to contribute to the current debate surrounding learning without awareness by triangulating different measures of awareness. Their objective was to determine the advantage and disadvantages of multiple measures of awareness (i.e., concurrent verbal reports, retrospective verbal reports, and subjective measures of awareness). The study confirmed that learners are able to rapidly acquire novel form-meaning connections under incidental learning conditions and without the benefit of feedback (Rebuschat, et al., 2015). In terms of the different measurements of awareness, the inclusion of the think-aloud procedure occasionally revealed that awareness had emerged earlier than participants reported in the interviews, indicating that retrospective recall can be unreliable, but useful in that they revealed partial rules that participants may have formed that could explain their performance. However, this could also be the cause of prompting participants to verbally describe rules or patterns at the end of the experiment. Additionally, the subjective measures of awareness allowed the detection of both implicit and explicit knowledge.

Dienes (2011) has advocated the use of subjective measurements in order to assess whether the knowledge acquired is conscious or unconscious. However, measurements of awareness are not typically featured in the statistical learning paradigm, but are more common in research about implicit learning (Hamrick & Rebuschat, 2011). This is as probably because the experiments within the paradigm usually are conducted on infants or children. There is, however, a possibility to administer basic measurements of awareness when the experiments are conducted on adults. One study by Hamrick and Rebuschat (2011) investigated whether the knowledge acquired in a typical statistical learning experiment was conscious, unconscious or both and concluded that statistical word learning in adults could result in both implicit and explicit knowledge. Rebuschat, et al. (2015) had similar results that indicated that the incidental exposure can results in both implicit and explicit knowledge of language.
Since language comprehension and production are thought to be partly based on implicit knowledge, it seems important to determine whether subjects in statistical learning research develop this type of knowledge. The lack of awareness is often assumed, but not empirically assessed (Aslin, Saffran, & Newport, 1999), leading to uncertainty regarding whether statistical learning typically results in conscious or unconscious knowledge (Hamrick & Rebuschat, 2011). The results from measures of awareness should however be interpreted with caution. Testing awareness can lead to biased results, for instance if a subject chooses to withhold conscious knowledge held with little confidence (Shanks, 2005). By incorporating a self-rating questionnaire as a measurement of awareness in the present experiment (See Appendix A), we got an idea of the degree to which the participants engaged in hypothesis testing or used other forms of explicit strategies during the experiment. This would also provide us with a deeper understanding of the knowledge acquired (Rebuschat & Williams, 2011).

**Individual Differences**

While researchers the last few decades have had an increasing interest in statistical learning, it is not until a few years ago that researchers have begun focusing on the interaction between individual differences and statistical learning and tried to demonstrate a correlation with other cognitive tasks (Arciuli & Torkildsen, 2012). The concept that people have different abilities for acquiring language is well established, but there is limited knowledge about the individual differences in equivalence of statistical learning (Brooks, Kwoka, & Kempe, 2016; Ryen, 2005). Individual differences can apply to differences in basic cognitive abilities like phonological short-term memory (Baddeley, Gathercole, & Papagno, 1998; Gupta & MacWhinney, 1997), statistical learning ability (Granena, 2013; Linck et al., 2013; Speciale, Ellis, & Bywater, 2004), nonverbal intelligence (Andringa, Olsthoorn, van
DISTRIBUTION OF EXPOSURE AND SUBCATEGORY LEARNING

Beuning, Schoonen, & Hulstijn, 2012; Grigorenko, Sternberg, & Erhrman, 2000) and 
verbal working memory capacity (Linck, Osthus, Koeth, & Bunting, 2014).

The difference between implicit and explicit learning has informed theories about the 
intellectual prowess of humans (Burns, Baetu, & Urry, 2016). The statistical learning research 
mostly focuses on infants or child language acquisition, though studies with adult subjects are 
also common (Rebuschat & Williams, 2011). Several studies even suggest that statistical 
learning is invariant with age (Gerken, et al., 2005; Gómez & Lakusta, 2004; Kittleson, 
Aguilar, Tokerud, Plante, & Asbjørnsen, 2010; Saffran, Aslin, et al., 1996; Torkildsen, et al., 
2013). A number of studies have shown that similarities in learning exist in adults, infants and 
children, suggesting these skills are present early in the development and continues into 
adolescence, for instance, infants, children and adults processes dependent probabilities so that 
they are able to detect words and tones in continuous speech (Saffran, 2002; Saffran, Aslin, et 
al., 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). They are also equally capable 
to detect these conditional probabilities in sequentially presented visual stimuli (Fiser & 
Aslin, 2002). This tells us how some similar factors are at play in how infants and adults 
detect isolated dependencies in sequential structures, also concerning statistical learning 
(Gómez, 2002). Based on the mentioned research demonstrating that infants (e.g., Gómez & 
Lakusta, 2004; Saffran, Aslin, et al., 1996), children (e.g. Plante et al., 2014) and adults (e.g. 
Kittleson, et al., 2010; Torkildsen, et al., 2013) have the same ability to learn statistical 
information we found it justifiable to use adult participants in the present study.

When learning takes place in adulthood, there is evidence to suggest that large 
individual differences in language learning abilities exist (e.g. Csizér & Dörnyei, 2005; 
Johnson & Newport, 1989). There has also been proposed that the individual variation in 
language acquisition abilities is linked to stable genetic differences (Ryen, 2005; Wong, 
Morgan-Short, Ettlinger, & Zheng, 2012). Both typically developing individuals and people
with learning disabilities have different language skills and basis for language acquisition, but there is less research on these individual differences in relation to implicit statistical learning (Ryen, 2005). The same kind of individual differences that can affect language acquisition may also influence statistical learning and language outcomes (Erickson & Thiessen, 2015). For instance, some aspects of the basal ganglia, hippocampus and prefrontal cortex’ functions are thought to influence statistical learning, which are influenced by the level of dopamine activity that is disparate among humans (Karuza et al., 2013; McNealy, Mazziotta, & Dapretto, 2006). Therefore, one can see how individual differences in statistical learning may partly affect the personal differences in language acquisition linked to differences in dopamine, and may consequentially have influences that differ on the process and outcomes of statistical learning (Erickson & Thiessen, 2015). More knowledge is still needed in relation to the neural mechanisms that support statistical learning and how these structures remains invariant or changes through development (Aslin & Newport, 2012).

A recent study by Brooks, et al. (2016) investigated the distributed effects and individual differences in second language (L2) learning on normally developed adults. Since structure of the input and learners’ nonverbal intelligence might affect the outcomes of learning, they wanted to see whether learning and generalization of L2 might be facilitated by less predictable input. The study also evaluated possible contributions of statistical learning ability, phonological short-term memory, verbal working memory and nonverbal intelligence in learning (Brooks, et al., 2016). Similar to the present experiment, the participants were exposed to Russian stimuli, but they had either a balanced or distorted item distribution in the input. They found that nonverbal intelligence and ability of statistical learning predicted learning, and also that nonverbal intelligence alone anticipated generalization of case-marking inflections to new language (Brooks, et al., 2016). Less predictable and more balanced input were related to more efficiency in generalization, but this was the case only for the first test
phase. The results from the study indicated that individual differences are more sustained in L2 acquisition than instructed teaching manipulations that vary the predictability of spoken elements in the input (Brooks, et al., 2016).

Several studies have indicated that language impaired individuals have poorly performance on statistical learning tasks, which is determined among children (Evans, Saffran, & Robe-Torres, 2009), adolescents (Hsu, Tomblin, & Christiansen, 2014; Tomblin, Mainela-Arnold, & Zhang, 2007) and adults (Grunow, Spaulding, Gómez, & Plante, 2006). Through these studies, one can see how the same underlying mechanisms seem to support the relationship between language learning abilities and statistical learning abilities (Romberg & Saffran, 2010). Arciuli and Torkildsen (2012) pointed out the need for longitudinal data to advance knowledge on the role statistical learning might play in the process of language acquisition, both in the context of normally developed and language impaired individuals. Longitudinal data could give more advanced knowledge on the field about how language acquisition might be influenced by statistical learning in normally developed and language impaired individuals.

Clinical Implications of Statistical Learning

To improve the quality of language treatment designs it is necessary to reduce the differences between theory and practice, and it is crucial that speech-language pathologists understand the principles of learning better (Alt, Meyers, & Ancharski, 2012). Within the field of speech-language pathology, studies demonstrating impaired statistical learning in individuals with language learning disabilities are of particular interest (Alt, et al., 2012). The general principles of learning from research on impaired learners can be used when designing language therapy interventions and exemplify how knowledge can be carried out in the clinical practice of speech language pathologists (Alt, et al., 2012). Insights from research from the implicit and statistical learning traditions might provide a better understanding of
language acquisition in children and adults and improve the methods for treating people with language difficulties (Arciuli & Torkildsen, 2012).

Research within the statistical learning paradigm is beginning to uncover what facilitates and what hinders learning in the context of language acquisition and language therapy. However, additional empirical research is still needed, especially with the hypothesized causal relationship between reduced statistical learning ability and language impairment in mind (Hsu & Bishop, 2011). The outcome of treatment depends on the effectiveness of the treatment exposure as well as the schedule of delivery and amount of treatment administered (Meyers-Denman & Plante, 2016). Clinical and experimental studies that have investigated the effect of distribution of exposure show different learning advantages when it comes to distribution of exposure (e.g. Barratt, et al., 1992; Childers & Tomasello, 2002; Meyers-Denman & Plante, 2016; Riches, et al., 2005; Smith-Lock et al., 2013; Ukrainetz, et al., 2009; Vlach, et al., 2008). Distributed exposure has been demonstrated to crucially affect word learning performance (Childers & Tomasello, 2002; Schwartz & Terrell, 1983). Moreover, children with SLI may be particularly sensitive to the effect of spacing, with reason to believe that distributed learning may be the most beneficial (Riches, et al., 2005). Furthermore, evidence suggests that distributed training is especially beneficial for poor learners, such as adults with dementia (Camp, Foss, O'Hanlon, & Stevens, 1996) children with learning disabilities (Gettinger, Bryant, & Mayne, 1982), and children with intellectual disabilities (Yoder, Woynaroski, Fey, & Warren, 2014). It has even been observed in a name-face matching task with non-impaired adults that poorer learners tend to benefit most from distributed practice (Cull, Shaughnessy, & Zechmeister, 1996). In addition, massed treatment has been reported to led to great benefits in expressive language outcomes (Barratt, et al., 1992). The studies mentioned above indicate that distributed and massed treatment can lead to exceptional learning effect. Depending on the goal of the treatment it
can also give clinicians some flexibility in terms of the dose schedule they use to deliver this
treatment in an evidence-based manner (Meyers-Denman & Plante, 2016).

With these findings in mind, two hypotheses exist to explore whether principles of
distribution of exposure might be beneficial in a language acquisition design. Findings from
studies on dose schedule may inform speech-language pathologists about distribution of
exposure when designing treatments for people with language difficulties (Meyers-Denman &
Plante, 2016). The participants in the study were adults with no self-reported language
impairments and we investigated the distribution of exposure principle in a laboratory setting.
Based on mentioned research demonstrating that infants, children and adults have the same
ability to learn statistical information (Gerken, et al., 2005; Gómez & Lakusta, 2004;
Kittleson, et al., 2010; Saffran, Aslin, et al., 1996; Torkildsen, et al., 2013), adults were used
to obtain useful detailed information about distribution of exposure as well as awareness.

**Research Questions and Hypotheses**

The purpose of the present experiment was to investigate whether the distribution of
exposure might facilitate or hinder the learning of noun gender subcategories in an unfamiliar,
natural language (Russian). This was an attempt to bridge the gap between theory and practice
and a result of trying to uncover the best approaches regarding treatment of language
impairments (Alt, et al., 2012). Participants were exposed to either a massed condition (i.e.,
the familiarization phase heard three times continuously before being exposed to the test
phase three times) or a distributed condition (i.e., tested three times directly after hearing the
stimuli material). Firstly, we hypothesized that a learning effect would occur, measured by
acceptance of grammatical and ungrammatical items. Secondly, if learning occurred, there
will be a difference in the distribution of exposure between the groups (massed
vs. distributed). Finally, if learning occurred, it would be interesting to investigate if the
participants were able to explicit express what they had learned. By incorporating a self-rating
questionnaire as a measurement of awareness, we sought to get an idea of the degree to which the participants engaged in hypothesis-testing or used other forms of explicit strategies.

**Methodology**

**Research Design**

There is a clear methodological distinction between quantitative research and qualitative research. Qualitative research designs collect in-depth knowledge, while quantitative research designs include experimental and nonexperimental studies (Polit & Beck, 2012). With the use of a quantitative research design in the present study, we were able to answer the research questions presented in our experiment. Quantitative research is orderly and structured in the gathering and analyzing of information relevant to research questions (Polit & Beck, 2012). This approach enables us to obtain quantitative information as well as generalizability and the findings are grounded in reality rather than our personal beliefs (Polit & Beck, 2012). We also included some of the participants’ qualitative reports to gain some insight.

In the present study, an experimental design was used to manipulate the independent variables (group condition, noun marking, age, level of awareness and previous language skills) as well as assess changes in the dependent variables (i.e., the participants’ responses, measured by acceptance rate and response time). As mentioned earlier, the participants in the present experiment were divided into two separate conditions and assigned to either a massed condition or a distributed condition. The two groups included the same amount of stimuli, but differed in terms of how the stimuli were presented. The stimuli were delivered either as three cycles consecutively of 64 items (192 items in total) before the test phase (massed condition) or divided in three cycles of 64 items each (also 192 in total) separated by test phases in between (distributed condition). Learning was operationalized as significantly higher acceptance rate of grammatical items compared to ungrammatical items. Our experiment
might be referred to as having a quasi-experimental design since the experiment could be said to violate some of the rules for a classic experimental design considering the participants were assigned to the two conditions by pseudo-randomization to ensure that equal numbers of males and females were allocated to each condition.

Participants

Forty graduate students (20 men and 20 women) were included as voluntary participants based on a signed consent. The participants included in the experiment were the ones that responded first and met the inclusion criteria. None of them had any knowledge of Russian or other languages similar to Russian (e.g., Slavic languages). In addition, participants who reported difficulties considering language learning in general, developmental or acquired neurological disorders, or had a known hearing loss were also excluded from the experiment beforehand.

The participants were primary Norwegian speakers where two participants’ native language was Swedish, one participant’s native language was German and one participant’s native language was Punjabi. The two Swedes were in separate groups, the German was in the distributed condition and the participant’s with Punjabi as her native language was in the massed condition. However, all the participants reported that they understood and spoke Norwegian fluently. While one previous study has found that prior language experience could influence learning (e.g. Lany, Gómez, & Gerken, 2007), Kittleson, et al. (2010) found no significant effects of language background. In their study, Kittleson, et al. (2010) investigated whether differences in the participants’ language background would have an impact. The participants in the study had different native language backgrounds and examined word segmentation in an unfamiliar language (Norwegian). The results indicated that a common strategy was used to separate words from a continuous speech stream in a foreign language regardless of their language background (Kittleson, et al., 2010). In the present
experiment the intention of asking about the participants’ language knowledge was to control for the influence of prior language experience in two ways; to exclude participants with knowledge of Russian or other Slavic languages, and to analyze whether the number of languages the participants speak correlates with their results.

We collected the required number of 40 participants in a short period of time and with low expenses, in accordance with the framework for the master’s project. It is common for researchers to work with samples rather than with populations because it is cost-effective to do so. With a representative population the external validity of the experiment may increase, thus the results could be generalized to a larger population and replicated with different participants (Cozby & Bates, 2012; Polit & Beck, 2012). Certain sampling procedures are less likely to result in biased samples than others, but a representative sample can never be guaranteed. In the present study the participants that responded first and met the inclusion criterion were included in the experiment. The advantage of this technique is that we were able to obtain participants without spending a great deal of money or time on selecting the sample (Cozby & Bates, 2012). The convenience sample as well as the criterions that the participants were not to have any knowledge of Russian or languages similar to Russian (e.g., Slavic languages) and not to have any self-reported hearing impairment, language disabilities or neurological disabilities, made the sample restricted. Previous experiments have acknowledged the relatively enduring capacity of implicit learning in individuals with different types of disorders, such as Parkinson’s disease (Witt, Nühsman, & Deuschl, 2002) and amnesia (Meulemans & Van der Linden, 2003). There are mixed evidence for language difficulties affecting implicit learning (for review see Alt, et al., 2012). By studying different populations valuable contributions to the field of implicit and statistical learning could be acquired. However, for convenience purposes, the present experiment focused on a normal population of adults with no developmental or acquired neurological disorders. To ensure
high internal validity when examining a phenomenon, researchers can with certainty make
assumptions of cause and effect if it is done under highly controlled conditions (Arciuli &
Torkildsen, 2012). The external validity becomes problematic when transferring an
assumption of statistical language acquisition from a laboratory to a natural language learning
setting. The transferability together with highly restricted sample weakens the external
validity and make it harder to generalize the findings to populations different from the one
studied and it must be done with caution (Cozby & Bates, 2012).

   The participants were given a monetary compensation of NOK 150 for taking part in
the experiment intended to cover any transportation costs etc. The money was given in cash or
through the app “Vipps by DNB” after the completion of the experiment.

To reduce the possible impact of mortality (e.g. a potential drop out) we had a few
extra participants that we could contact if necessary. Mortality is the threat that arises from
attrition in groups being compared that could affect the internal validity of the study (Cozby
& Bates, 2012). None of the participants in the present study chose to withdraw. A reason for
this could be because the experiment did not collect data over time (Polit & Beck, 2012).
Even though it has been pointed out that longitudinal studies are needed to advance the
understanding of the relationship between statistical learning and language acquisition (Polit
& Beck, 2012), the use of studies over a longer period of time could be damaging for the
mortality.

Material and Procedure

   Every research design has shortcomings and it is important that the researcher is aware
of such possible limitations. In order for our research design to be accepted as scientific, a
high degree of validity and reliability is required. The research design should thus measure
what it is intended to measure as well as obtaining the same results during repeated testing
Pilot study. During the fall of 2016, a pilot study was carried out with six Norwegian graduate students. After the pilot study, all participants in the experiment were recruited and tested. The pilot study allowed us as researchers to complete a trial run with a small number of participants (Cozby & Bates, 2012) and was conducted to get an idea of whether the procedures and material were accurate and beneficial, whether the given instructions were effective and whether the experimental setting seemed reasonable. In addition to give conductive feedback, the pilot work allowed us to become more comfortable in the role as researchers, making it an educational experience, and standardize the experiments’ procedures. It is crucial that the researcher is aware of his/her relation to the participants in the experiment. By practicing how we as researchers were going to behave consistently during the experiment we were able to minimize the experimenter bias or expectancy effects (Cozby & Bates, 2012).

Recruitment. The participants in the present experiment were recruited through posters on message boards at some of the University of Bergen faculties combined with social media. The participants included were from the University of Bergen and the Bergen University College. The ability to replicate a study is an important concept for external validity and can be a way of overcoming any problems of generalization that occur (Polit & Beck, 2012). While using students highly restricts the population it is also a safeguard against the limited external validity of a single study, since it opens up for replication at other universities. Moreover, students are increasingly diverse and increasingly representative of the society as a whole (Cozby & Bates, 2012). However, we cannot know if the participants’ academic backgrounds are comparable.

Material. The present experiment was based on previous studies by Richardson, et al. (2006), Gerken, et al. (2005) and Eidsvåg, et al. (2015). The Russian auditory stimuli in the present experiment was partly obtained through Eidsvåg, et al. (2015) where 15 root words
were included from their study. Dr. Elena Plante at the University of Arizona provided the additional 29 root words. All the root words included had double-marked suffixes. A male and female native Russian speakers recorded the stimuli for the familiarization phase. All the test words were recorded by a female native Russian speaker and were different from the familiarization phase.

The use of natural language (Russian) is comparatively a new design so there is little evidence to acquire on what effect sizes that can be anticipated (Cohen, 1992). On the basis of Torkildsen, et al. (2013)’s and Eidsvåg, et al. (2015)’s sample size, the present experiment set a temporarily sample size of 40 participants with 20 participants in each of the two conditions (massed and distributed). An alternative approach to select the sample size is on the basis of a desired probability of correctly rejecting the null hypothesis ($H_0$) (Cozby & Bates, 2012) and depends on the sample size, value of alpha and the effect size (Cozby & Bates, 2012; Kirkwood & Sterne, 2003). Therefore, after the data was collected for the 40 participants, the power and alpha of the experiment was controlled to ensure that power was more than 80% and the alpha was 5%. The calculations indicated that even a study with almost 1000 participants would give similar results. Since the calculations gave these results the temporary sample size of 40 was maintained. However, if we had included a more robust sample, the experiment would have had an even more representative sample of the population aimed to describe (Cozby & Bates, 2012).

As mentioned, the participants in the present experiment were exposed to double-marked nouns with adjacent dependency through Russian auditory stimuli. Newport and Aslin (2004) have suggested that statistical learning of nonadjacent dependencies are more difficult to acquire than adjacent ones. The participants had no previous knowledge with the semantic context of the words they were presented. As a result, they had to rely on diminutive cues
(i.e., suffixes) and the relationship between the two alternative markings (i.e., \(-telyem\) and \(-telya\) vs \(-koj\) and \(-ku\)).

The Russian noun gender system includes three subcategories (masculine, feminine, neuter) and numerous suffixes (Corbett, 1991). The fact that the Norwegian language marks three subcategories of noun gender (masculine, feminine and neuter) could be of significance as it is similar to the Russian language subcategories. This could affect the relative ease in acquiring knowledge about a gender marking system. Conversely, English, for instance, does not have a noun gender system. This could potentially be a challenge when comparing results across languages that have grammatical markers and languages that do not have it. As earlier mentioned, the effects of statistical language learning have in many studies been explored with the use of a variety of artificial languages (Stanovich, 2010), which provides a high degree of experimental control. Using natural language instead of artificial language has its benefits on the account of artificial language not including the great variety and complexity found in natural language, as well as being greatly simplified. In addition, the use of a natural language can increase the external and ecological validity of the present experiment.

In the present study only the subsets of masculine and feminine double-marked gender markings and some possible suffixes were included. The exclusion of the one noun gender subcategory and other gender markings was to limit the tasks degree of difficulty and the use of double-marked words was to increase the variability of the words. Previous studies have shown the benefits of double-marked nouns in infants (Gerken, et al., 2005) and adults with a history of language-based learning disabilities (Richardson, et al., 2006) when it comes to generalization of gender marking "rules". In Eidsvåg, et al.’s (2015) experiment the participants showed a learning advantage with double-marked nouns, even though they were exposed to slightly more single-marked words due to a coding error. Consequently, each root word presented in the present experiment had two possible gender markings. There were two
variations of feminine double markings (i.e., -koj or -ku) and two masculine double markings (i.e., -tel + ya or -tel+yem). Any word that could take one feminine ending (i.e., Blondin + -koj) could also take the other feminine ending (i.e., Blondin + -ku). Similar, the masculine root word suffix –telya could be exchanged with the alternative suffix -telyem (i.e., Rastochi + telya; Rastochi + telyem).

Further, the stimuli were presented in a randomized order. This means that the root word with the different suffix could come sequential (e.g., Pisatelya and Pisatelyem). The intention of a randomized order was to highlight the way one naturally could come across words when learning an unfamiliar, natural language. To reduce this risk, we could have chosen to present the material in a block-randomized order. However, when going through the data set after all the participants, there were no root words that came in a sequential order that emphasized the relationship between the two possible suffixes.

**Familiarization phase.** The stimuli material consisted of 32 different root words with two suffixes each (64 unique items), listed in Table 1. The stimuli was delivered either as three cycles consecutively of 64 items (192 items in total) before the test phase (massed condition) or divided in three cycles of 64 items separated by test phases in between (distributed condition). Each cycles had an individual new random order. All participants heard the same amount of input, and were tested with the same amount of words, but with a different degree of distribution of exposure. Note that despite the differences in stimuli presentation that the participants in both conditions heard the same total number during the experiment. As mentioned, previous studies have shown that variability promotes learning (e.g. Eidsvåg, et al., 2015; Grunow, et al., 2006; Plante, et al., 2014; Torkildsen, et al., 2013).

In addition to the variability through the double-marked words (see Eidsvåg, et al., 2015; Gerken, et al., 2005; Torkildsen, et al., 2013), the auditory material for the present experiment was recorded by two different talkers (one male and one female) who were native speakers of
Russian. The sound files were edited to correspond to the actual length of the word and to produce approximately equal loudness across words. In addition, there was a 300-580 milliseconds break between each word.

Table 1

*Stimuli Set for the Familiarization Phase*

<table>
<thead>
<tr>
<th></th>
<th>Masculine words</th>
<th>Feminine words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female voice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dviga -telya/telyem*</td>
<td>Blondin -koj/ku*</td>
<td></td>
</tr>
<tr>
<td>Khrani –telya/telyem</td>
<td>Kodoerov –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Muchi-telya/telyem*</td>
<td>Progul –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Osnova-telya/telyem*</td>
<td>Rozoch –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Sluzhi –telya/telyem</td>
<td>Rubash –koj/ku*</td>
<td></td>
</tr>
<tr>
<td>Smotri –telya/telyem</td>
<td>Skovorod-koj/ku*</td>
<td></td>
</tr>
<tr>
<td>Sozda –telya/telyem</td>
<td>Vetrov –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Uchi –telya/telyem</td>
<td>Vystava –koj/ku</td>
<td></td>
</tr>
<tr>
<td><strong>Male voice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blagode - telya/telyem*</td>
<td>Babush –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Dushi -telya/telyem*</td>
<td>Devoch –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Grabi –telya/telyem</td>
<td>Juboch –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Pisa –telya/telyem</td>
<td>Karmel-koj/ku*</td>
<td></td>
</tr>
<tr>
<td>Potrebi –telya/telyem</td>
<td>Khlopush –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Rastochi –telya/telyem</td>
<td>Makush -koj/ku*</td>
<td></td>
</tr>
<tr>
<td>Razrushi –telya/telyem</td>
<td>Podush –koj/ku</td>
<td></td>
</tr>
<tr>
<td>Vodi -telya/telyem</td>
<td>Petrush-koj/ku</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The words were recorded equally by two native speakers of Russian, a male speaker and a female speaker. * indicates stimulus previously used in Eidsvåg et al. (2015).

*Test Phase.* The stimuli set for the test phase consisted of 12 different root words with two possible suffixes (24 unique items), equally split between masculine and feminine form
listed in Table 2. Half of the items were grammatically correct (i.e., grammatical items), while the other half violated the grammatical pattern (i.e., ungrammatical items) that the participants had been exposed to during the familiarization phase. In total, the participants in both groups heard 72 items. The stimuli was delivered either as three blocks consecutively with individual new random order of 24 items (massed condition) or divided into three blocks of 24 items separated by the familiarization phases (distributed condition). Each cycles had an individual new random order. Participants were asked to judge each item as a grammatically correct (press “smiley face”) or incorrect (press “frowny face”) Russian word in order to proceed to the next test item. The root words had two legal markings that were presented in the familiarization phase (e.g., feminine word: Vetrov –koj/ku, Masculine word: Potrebi -telya/telyem). To perform well on the test phase, the participants had to detect cues from the familiarization input. In the test phase each root word had either a grammatical correct ending (e.g., Deja –telyem) or an ungrammatical word ending (e.g., Deja – telu). Hearing any other ending than the four suffixes from the familiarization phase should have led the participants to reject the test item. A female native speaker of Russian recorded the words. This was a different woman’s voice than was heard in the familiarization, and was only heard during the test phase.
Table 2

**Stimuli Set for the Test Phase**

<table>
<thead>
<tr>
<th></th>
<th>Masculine words</th>
<th>Feminine words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grammatical/correct words</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deja –telyem</td>
<td>Brjunet –koj*</td>
<td></td>
</tr>
<tr>
<td>Osvezhi – telyem*</td>
<td>Obert –koj</td>
<td></td>
</tr>
<tr>
<td>Pokupa –telyem*</td>
<td>Tarel –koj</td>
<td></td>
</tr>
<tr>
<td>Rodi –telya</td>
<td>Devush –ku*</td>
<td></td>
</tr>
<tr>
<td>Stroi –telya</td>
<td>Maka –ku*</td>
<td></td>
</tr>
<tr>
<td>Ljubi – telya*</td>
<td>Verev –ku</td>
<td></td>
</tr>
</tbody>
</table>

|                |                  |                |
| **Ungrammatical/incorrect words** |                |                |
| Deja –telu     | Brjunet -kya     |                |
| Osvezhi –telyem* | Obert –kya   |                |
| Pokupa –telyem* | Tarel –kya      |                |
| Rodi –teloj    | Devush –kyem*   |                |
| Stroi –teloj   | Maka –kyem*     |                |
| Ljubi –teloj*  | Verev –kyem     |                |

*Note.* The words were recorded by a different female Russian speaker than in the familiarization phase. * indicates stimulus previously used in Eidsvåg et al. (2015).

**Self-rating Questionnaires.** After completing three cycles of familiarization phase and test phase, the participants were asked to fill out a self-rating questionnaire. The questionnaire was previously used by Eidsvåg, et al. (2015), however, in this study, minor changes were done to the questionnaire. The wording of some questions was changed. By changing the wording of the previous question about regularities, we tried to increase the reliability so that the question was easier to interpret and less ambiguous. The main goal of the questionnaire was to gather information about the participants’ understandings and impressions after completing the experiment. Information concerning the participants’ age and number of languages known by the participants was also collected. In addition, questions with increased
specificity were added, since we wanted to identify the participants’ awareness on a range of awareness. From the more general questions about their recall of what they had been through, we then asked more specific questions about their awareness for the suffixes they had encountered. However, it is important to note that the participants were exposed to auditory stimuli in the experiment and were later presented with the Russian suffixes in a written manner in the questionnaire.

When it comes to the distributed condition, the participants could adjust and modify their expectations to the experiment and applied explicit strategies in compliance with their hypothesis or in accordance with what they believed was the desired behavior (c.f., Hawthorne effect, acquiescence bias, demand characteristics etc.) (Polit & Beck, 2012).

However, with the answers from the self-reported questionnaire, we were able to deduce that the level of awareness of both conditions was the same. One of the more specific questions in the questionnaire asked the participants to give six suffixes (three grammatical and three ungrammatical) a score from 0-100 by their thought of probability that a suffix occurred during the experiment. The mean for the grammatical correct items were made a variable. This applied for the ungrammatical items as well. The questionnaire originally had eight suffixes the participants had to judge the probability of occurring. We only included six suffixes to have a better basis for comparison. This could, however, have distorted the data by us choosing the suffixes to be excluded. In addition, the other questions that were meant to be more specific were excluded, since they appeared to not be good subjective measurements, as they did not provide sufficient information about their awareness.

Procedure. The participants were tested individually in a research lab at the University of Bergen. They completed the experimental tasks and then filled out a self-reported questionnaire after signing an informed consent form (see Appendices A and B). The participants were seated in front of a laptop, asked to put on headphones, and were given
instructions on how to adjust the volume. To improve the comparability across participants, the same information and instructions were given to all participants. The researcher left the room, but remained at hand in the event of any technical issues occurring. The two researchers equally divided the workload and the subjects were assigned to a researcher by availability. Prior to each familiarization phase and test phase, the participants were given written instructions on the computer screen (Appendix C). The information was to make sure that the participants to a certain degree would know what was going on. The instructions were carefully composed to avoid that the participants received more information than needed. We believed that the participants already got some expectations about the experiment through the test situation. However, information as to the differences in distribution of exposure (massed vs. distributed) and what they should notice about the stimuli (i.e., subcategory gender markings for double, feminine/masculine words) was not given. The experiment lasted for about 15 minutes, with the additional questionnaire given out afterward.

**Data Processing and Analyses**

Stimuli presentation and data collection were performed with E-Prime 2.0 Professional (Schneider, Eschmann, & Zuccolotto, 2002) and processed and analyzed with the use of Excel (Version 15) and in Statistica (Version 13). The data was statistically analyzed with descriptive statistics, mixed between-within analyses of variance, post hoc tests and correlational analyses. One of the most important interpretive tasks is to assess whether the results from the analyses are right. If the results are not credible, the remaining interpretive issues (meaning, magnitude, precision, generalizability) are not likely to be relevant (Polit & Beck, 2012). A study can be considered to have high internal validity if with considerable certainty one variable can presume the effect of another variable (Polit & Beck, 2012). With a description of the experiment in the method section, the present study will make the
experiment transparent, making other researchers able to replicate and ensure that the findings are not due to errors or biases by us as researchers (Cozby & Bates, 2012).

The variables in our study had different measuring scales. Age, response time and acceptance of grammatical and ungrammatical items were measured on ratio scales since the measurements provide information about critical attribute, the intervals between the participants and because they are rational. The participants’ degree of awareness was measured on an ordinal scale, which is because the measurement captures information about not only equivalence, but also about relative rank. The group condition and participants’ sex contains categorical information and are within the nominal scale (Polit & Beck, 2012).

In the present study we needed to make the dependent variables for the learning condition (i.e. items accepted as correct and response time for test items) measureable to ensure the construct validity. A variable can be empirically studied by operationalizing the concept (e.g., Gómez & Gerken, 2000; Torkildsen, et al., 2013). The acceptance rate for the test items and the participants’ response time acted as the dependent variable. Acceptance rate is typically utilized in experiments similar to the present study in which learning is defined as significantly higher acceptance rate of grammatical items compared to ungrammatical items. Learning is then defined as significantly higher accept of grammatical items compared to ungrammatical items. This method is based on principles from the signal detection theory and contributes to control for an underlying bias towards accepting or rejecting items overall (MacMillan & Creelman, 2005). The participants’ answers from the test phase are sorted into one of four categories as shown in Table 3. An alternative method would be to use accuracy rate (i.e., calculation of the total number of correct answers), however, acceptance rate is thought to be a more precise measurement (Hochhaus, 1972).
Table 3

The Four Categories for Stimuli and Response

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Response: Yes (accept)</th>
<th>Response: No (reject)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical items</td>
<td>HIT</td>
<td>MISS</td>
</tr>
<tr>
<td>Ungrammatical items</td>
<td>FALSE ALARM</td>
<td>CORRECT REJECTION</td>
</tr>
</tbody>
</table>

Note. Learning was defined as the significantly higher acceptance for grammatical items ("HIT") compared to ungrammatical items ("FALSE ALARM"). Adapted from “Detection Theory: A user’s manual” by N. A. MacMillan & C. D. Creelman, 2008, published in Mahwah, New Jersey by Lawrence Erlbaum Associates.

The main focus of the experiment was the participants’ acceptance rate, however we also measured the participants’ response time. Other factors could affect learning, therefore the present experiment will examine the interaction between the participants’ acceptance of grammatical and ungrammatical items and their response time. It has for a long period of time been well known that one can tradeoff accuracy for speed in supposedly any task (Wickelgren, 1977). Speed-accuracy tradeoff (SAT) is a dilemma which occurs when participants are making decisions in different situations and need to decide between the competing demands of acceptance rate and response time (Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010). This relationship is often described as the willingness to respond slowly and make relatively fewer mistakes compared to responding quickly and making more errors (Zimmerman, 2011). This behavioral tendency for response time to interact with acceptance rate seems to be inevitable when it comes to forced choice between two alternatives, and SAT studies are mostly favored to experiments about response time of the actions of memory, perceptual and other cognitive processes (Heitz, 2014; Webb & Adler, 2008).
In the data set there was some missing data. By excluding cases pairwise we made sure that participants were included in any of the analyses for which they had necessary information. As a consequence, no participant was totally excluded from the analyses that which would have limited your sample size (Pallant, 2013).

**Analysis of variance.** An analysis of variance (ANOVA) is used when you need to control for an additional variable that may be influencing the relationship between the independent and dependent variable (Polit & Beck, 2012). The observed relation between the independent and independent variable in the present experiment could have been affected by variables not accounted for (i.e., room condition, information given to the participants, time of day). This can potentially influence the internal validity of the experiment (Polit & Beck, 2012). However, we tried to keep the variables not tested constant to manage for uncontrolled third variables.

To test our hypotheses, we needed to combine a between-subjects design with a within-subjects design. Through the parametric test, mixed between-within ANOVA, we were able to combine the “within” analyses (repeated measurements, acceptance rate and response time) and the “between” analyses (group condition) in one process (Greer & Mullhern, 2011; Pallant, 2013). The independent variables (group condition, repeated measures) could thus be seen in the context of the behavioral dimensions (acceptance rate and response time). The design also enabled us to analyze for interaction effects to see if the responses developed differently over cycles under the two conditions.

In the present study we wanted to examine possible effect of distribution of exposure through having the differing conditions (massed condition, \(n = 20\); distributed condition, \(n = 20\)), and to assess if results were clinically important. In order to see whether there was a change over the repeated exposures in the group of all participants (= 40), repeated measurements were used to explore possible differences between measurement point one, two
and three (Pallant, 2013). For the ANOVA, the measure of the effect size is included in additional to the $F$-, $t$- and $p$-values in order to enable comparisons between different analyses on a common measurement scale. Here the effect size it measured by calculating the partial eta-squared ($\eta_p^2$), which includes the estimate of the within effect, combined with all other possible effects. The partial eta-squared calculated in the ANOVA is a measure of variance that includes the estimate of the within effect combined with all other possible effects.

**Post hoc tests.** In general, post hoc comparisons are used to evaluate all potential pairs of groups following a significant test of overall group variations (e.g., in an ANOVA) (Polit & Beck, 2012). In order to explore the statistically significant effect of acceptance rate (accepted grammatical vs. ungrammatical items) and group condition (massed versus distributed), the present study conducted a Tukey’s HSD. The use of post hoc tests made it easier to isolate the differences between group means that are responsible for rejecting the overall ANOVA null-hypothesis. It would be possible to use a series of t-tests, but this would have increased the possibility of a Type I error (Polit & Beck, 2012).

**Correlation.** Correlational analyses were conducted to examine the relationship between the measures of acceptance rate, response time and responses on the self-rating questionnaire and the participants' ‘response. In the present experiment we hypothesized that a correlation between the difference in distribution of exposure (massed vs. distributed) and degree of learning measured by acceptance rate (the difference between grammatical and ungrammatical items) would appear. The participants’ answers in the self-rating questionnaire about degree of awareness, rule searching, randomness and assertiveness during testing were quantified and graded from 0 to 3 (0 = not aware/random to a very large extent/did not follow rules, 3 = aware/not random/followed rules). In the present experiment parametric tests were used to analyze variables concerning acceptance rate, response time and the responses on the questionnaire. Correlation procedures are often used when describing relationships between
two variables (Polit & Beck, 2012). When high values on one variable are associated with high values on a second variable, a positive correlation occurs. A negative relationship, however, indicates that as one variable increases, the other decreases (Polit & Beck, 2012). A perfect correlation, either positive or negative, signifies that the value of one variable can be ascertained precisely by knowing the value of the other variable. There is also a possibility of there being no relation between the variables (Pallant, 2013).

**Ethical Considerations**

Ethical considerations are important aspects when contemplating the quality of a study (e.g. Eidsvåg, et al., 2015; Plante et al., 2015; Torkildsen, et al., 2013). It is essential for a researcher to recognize the impossibility of total objectivity, though there is also a desire to understand and represent the real-world in an experiment (Polit & Beck, 2012). We attempted to be as neutral as possible by being aware of the ethical and fundamental research issues (Polit & Beck, 2012). With the research issues discussed in the method section and aspects that will be presented, we believe the study can be considered ethically justifiable and of no risk to the individuals involved. The participants’ interest and needs were attended to with care and the suitable tests were applied in a satisfactory environmental setting.

By being aware of the researcher's responsibility we were able to optimize benefits and make the study as little harmful as possible. Participants in the present experiment were not exposed to damage or discomfort, whether physical, psychological or economical (Polit & Beck, 2012). The experiment lasted for approximately 15 minutes with a survey at the end that was not too strenuous. As for the economical discomfort, the participants were given 150 NOK (approximately 18 USD) as a compensation for participating in the study. The amount was set to compensate for any expenses in relation to the experiment. With the payment not being too large, it served as a reasonable compensation for the participants’ time and not as a way to “buy” participation.
The participants were informed that participation was voluntary and that the data would be anonymized. This was to assure and protect the participants and also give them adequate information about the experiment (Cozby & Bates, 2012). Information about the experiment was sent out to the participants, and handed out again before the experiment took place. It included the experiment’s purpose, contact information to the researchers and supervisors, information about anonymity, safekeeping of information, and the possibility to withdraw from the study with no loss of benefits (Helsinki, 2013; Polit & Beck, 2012) (see Appendix D). This was to make sure that the participants knew that withdrawing from the study would have no consequences and that there would be no judgmental treatment and they were thence able to voluntarily consent or decline participation. All essential information was given to the participants prior to signing the informed consent form and participation. To protect the participants’ right to privacy any sensitive information was deleted or destroyed after the completion of the study.

All research studies that involve personal information that can be traced directly or indirectly should be reported to the Norwegian Center for Research Data (Norsk senter for forskningsdata, NSD). In the present experiment, no personal information that could be connected to the data and statistical analyzes were collected. Since we only handled anonymous information there was no obligation to report the project and get approval from NSD. In addition, the present study was not related to health sciences, so it did not come underneath the authority area of Regional Committee for Medical and Health Research Ethics (Regionale komiteer for medisinsk og helsefaglig forskningsetikk, 2012 REK). The participants were given identification numbers from 10 to 49, where age and gender were admitted, so the only identification we had that connected the data and survey answers were the numbers appointed to the participants that had no link to any information that could identify the participants.
While withholding information might be unwise, providing too much information could potentially invalidate the results of the study (Cozby & Bates, 2012). In advance of the study the participants were only given information that allowed them to get a basic idea of what the intention of the study was, subsequently, omission of information was involved. In the present study the participants were aware that they participated in a study on statistical learning, but did not know the hypothesis about massed and distributed exposure.

**Concluding Remarks**

The present experiment investigated whether distribution of exposure would affect initial learning of noun gender subcategories in an unfamiliar, natural language by dividing the participants into the two conditions massed and distributed. Furthermore, by measuring the participants’ response time, we were able to get an indication of a possible speed accuracy trade off. In addition to these objectives, we sought to get an idea of the degree to which the participants engaged in hypothesis-testing or used other forms of explicit strategies during the experiment. This was examined by incorporating a self-report questionnaire form that the participants filled out after the experiment.

The findings revealed that participants in the two different conditions (massed and distributed) had an equal amount of benefits for learning and did not treat grammatical and ungrammatical items differently at a statistically significant level. While both conditions started to accept more grammatical items throughout the experiment, they were also inclined to accept more ungrammatical items. These results provide support for the two distribution of exposure conditions. In addition, the measurement of awareness revealed that the participants had similar awareness during the experiment while also reporting an inclination to accept grammatical items more than ungrammatical items.

The main findings of our experiment contributes to the debate around distribution of exposure in language acquisition. Previous studies have demonstrated distributed learning’s
beneficial effect on learning (Childers & Tomasello, 2002; Janiszewski, et al., 2003; Riches, et al., 2005; Smith-Lock, et al., 2013; Vlach, et al., 2008), but findings in favor of massed exposure have also been found (Barratt, et al., 1992). Similar to our experiment, Meyers-Denman and Plante (2016), Ukrainetz, et al. (2009) and Proctor-Williams and Fey (2007) both types of distribution of exposure had a learning effect. With the present experiment we have extended the evidence base by reporting the different effects associated with distribution of learning noun gender subcategories. The findings may enlighten researchers interested in the role of distribution of exposure in statistical learning experiments. The subjective measure of awareness could provide us with a deeper understanding of the knowledge acquired in statistical learning experiments as well as the nature of the resulting knowledge.
References


https://helseforskning.etikkom.no/reglerogrutiner/soknadsplikt?p_dim=34997


DISTRIBUTION OF EXPOSURE AND SUBCATEGORY LEARNING


DISTRIBUTION OF EXPOSURE AND SUBCATEGORY LEARNING


The Effect of Distribution of Exposure on Unguided Subcategory Learning in Adults

Kristina T. Osebakken and Iselin A. C. Partee

Master’s Programme in Health Sciences, Logopedics

The Faculty of Psychology
Department for Biological and Medical Psychology

University of Bergen

Spring 2017
Abstract

The present experiment investigated whether distribution of exposure would affect initial learning of a noun gender marking system in an unfamiliar, natural language (Russian). Forty graduate students (20 men and 20 women) participated in an experiment where they were familiarized and tested for learning gender marking of Russian nouns. The stimuli were delivered either as three cycles consecutively of 64 items (in total 192 items) before the test phase (massed condition) or divided into three cycles of 64 items separated by a test phase (distributed condition). Both groups heard the same total of 192 items throughout the experiment. Learning was operationalized as significantly higher acceptance rate of grammatical items compared to ungrammatical items. Additionally, the participants filled out a questionnaire try to measure their degree of awareness during the experiment. The findings revealed that both groups showed evidence of learning as measured by higher acceptance of grammatical items than ungrammatical items. However, the participants’ response time and acceptance rate could indicate that the two conditions triggered different learning strategies. The questionnaire results revealed that the conditions had similar degree of awareness and were more inclined to accept grammatical items than ungrammatical items.

Keywords. Statistical learning, distribution of exposure, natural language, awareness
The Effect of Distribution of Exposure on Unguided Subcategory in Adults

The field of language acquisition has been greatly influenced by discoveries of infant and adult learners ability to gather information from speech or constant input and using this information to estimate distinct structures (Bulgarelli & Weiss, 2016; Saffran, Aslin, & Newport, 1996). Research conducted on statistical learning, i.e., being able to use statistical information in the environment when acquiring a language, has had increasing interest and seems to be profitable (Rebuschat & Williams, 2011). Earlier research on statistical learning (see Saffran, Aslin, et al., 1996; Saffran, Newport, & Aslin, 1996) has been important by demonstrating that infants are able to make extensive use of environmental cues when acquiring language (Ellis, 2006). The distribution of exposure in treatment could be a possible reason for this ability (Meyers-Denman & Plante, 2016).

The aim of the present experiment was to investigate whether distribution of exposure might facilitate learning of noun gender subcategories in an unfamiliar language. Within language learning studies, distribution of exposure is usually distinguished by the deliverance of doses. It can either be delivered clustered within a compressed period of time (massed condition) or spaced across a longer period of time (distributed condition) (Meyers-Denman & Plante, 2016). A number of studies have reported that differences in exposure should be taken into account when dealing with language acquisition and have identified distribution of exposure as an important element that may affect the efficacy of treatment (e.g., Barratt, Littlejohns, & Thompson, 1992; Childers & Tomasello, 2002; Meyers-Denman & Plante, 2016; Proctor-Williams & Fey, 2007; Riches, Tomasello, & Conti Ramsden, 2005; Smith-Lock et al., 2013; Ukrainetz, Ross, & Harm, 2009; Vlach, Sandhofer, & Kornell, 2008). Research concerning massed versus distributed learning suggest that enhanced learning follows distributed practice (e.g., Janiszewski, Noel, & Sawyer, 2003). In a meta-analysis of 97 studies distributed learning was proven to be more beneficial for an extensive variety of
linguistic objectives given in numerous modalities to adults and children (Janiszewski, et al., 2003).

Other studies have suggested the benefit of distributed learning in children (Riches, et al., 2005; Smith-Lock, et al., 2013; Vlach, et al., 2008). The effect of distribution of exposure has been investigated by Smith-Lock, et al. (2013) in one study concerning language treatment. This study consisted of a grammar treatment program on 5-year old children with specific language impairment. The massed condition consisted of treatment administered one time per day for 8 days in 1-hour sessions, while the distributed conditions had treatment one day per week for 8 weeks, also in 1-hour sessions. The children that received treatment over 8 weeks, showed a significantly increment in scores compared to the children in the 8-days group. Additionally, evidence suggest that distributed treatment is especially beneficial for poor learners, such as adults with dementia (Camp, Foss, O'Hanlon, & Stevens, 1996), children with intellectual disabilities (Yoder, Woynaroski, Fey, & Warren, 2014) and children with learning disabilities (Gettinger, Bryant, & Mayne, 1982). Researchers have suggested that the benefit of distributed learning could be attributed to factors such as greater variability of contextual cues during encoding (Glenberg, 1979) and the challenge for memory retrieval when items are distributed over a period of time (Schmidt & Bjork, 1992).

The results from the abovementioned studies argue in favor of distributed treatment. Conversely, one study demonstrated that a more intensive treatment schedule (massed) led to greater progress in expressive language outcomes compared with less intensive interventions (Barratt, et al., 1992). 2- to 5-year-old children had an overall treatment duration controlled within approximately 6 months. The treatment was administered either four times a week in a 3-week period within each of two 3-month cycles or one time weekly for 6 months. The researchers concluded that the massed treatment led to greater benefits than the distributed
treatment, even though the study lacked strong experimental control over distribution of
dosage and related distribution of treatment sessions (Barratt, et al., 1992).

Furthermore, other studies have examined the effect of distribution of exposure for
treatment and have found conflicting results (e.g., Meyers-Denman & Plante, 2016;
Ukrainetz, et al., 2009). For instance, the effects of two different intervention schedules was
examined by Ukrainetz, et al. (2009). They examined these two frequency dosages of
treatment on at-risk kindergartners’ acquisition of early literacy knowledge concerning
phonemic awareness. The findings from this study indicated no difference in performance
between the two groups, except for one subskill that favored the individuals who were trained
over a longer period of time(Ukrainetz, et al., 2009). However, the main conclusion was that
distribution of exposure did not have any considerable impact on the results, demonstrating
that weekly continuous treatment were similar to those that had a short and intensive
treatment(Ukrainetz, et al., 2009). Meyers-Denman and Plante (2016) got similar results when
they examined the effect of distribution of exposure on treatment of grammatical morphology
deficits in children with specific language impairment. The findings indicated an overall
treatment efficacy, but showed no differences in treatment effects from the massed and
distributed condition (Meyers-Denman & Plante, 2016).

Both experimental and treatment studies have such contrasting results when it comes
to massed versus distributed learning suggesting that either one could lead to a greater
treatment effect (Meyers-Denman & Plante, 2016). In the present study we sought to
investigate whether the benefit of distributed learning found in learning grammatical forms
(Vlach, et al., 2008), lexical learning (Riches, et al., 2005), and language treatment (Smith-
Lock, et al., 2013) could be transferred to learning the gender marked system in an unfamiliar
language. The present study might be the first to look at the interaction between the statistical
learning and distribution of exposure when it comes to noun gender subcategories.
In order to account for other factors or cognitive processes affecting learning, the present experiment examined the interaction between the participants’ acceptance of grammatical and ungrammatical items and their response time. A dilemma known as the speed-accuracy tradeoff (SAT) is when participants are making decisions in different situations and need to decide between the competing demands of acceptance rate and response time (Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010). This relationship is often described as the willingness to respond slowly and make relatively fewer mistakes compared to responding quickly and therefore making more errors (Zimmerman, 2011). This behavioral tendency for response time to interact with acceptance rate seems to be inevitable when it comes to forced choice between two alternatives (Heitz, 2014).

When it comes to the method used to investigate in statistical learning research the way input is carefully manipulated might be one of the most distinctive features (for a review, see Gómez & Gerken, 2000). The input created for these studies typically establish the defining feature of an artificial language (e.g. Gómez, 2002; Gómez & Lakusta, 2004; Mintz, 2002; Saffran, Aslin, et al., 1996). Studies concerning massed versus distributed learning have involved various forms of both artificial and natural language learning in which materials were presented and tested verbally (Meyers-Denman & Plante, 2016). The use of artificial language stimuli creates a challenge to ecological validity of the studies (Erickson & Thiessen, 2015; Romberg & Saffran, 2010; Saffran, Aslin, et al., 1996). In language studies using artificial stimuli, learning is facilitated by the presence of multiple cues (Gómez & Lakusta, 2004; Morgan, Meier, & Newport, 1987; Thiessen, Hill, & Saffran, 2005). Still, it is unknown how these certain learning mechanisms operate when it comes to natural language.

Similar to Romberg and Saffran (2010), the present study aims to address the value of reducing the gap between statistical learning and real-world learning. Moreover, to avoid
concerns about ecological validity, the present study wanted to test the statistic language learning process by using natural stimuli.

The stimuli material in the present experiment consisted of a subset of Russian masculine and feminine nouns combined with double-marked suffixes indicating the grammatical gender. Gerken, Wilson, and Lewis (2005), Richardson, Harris, Plante, and Gerken (2006) and Eidsvåg, Austad, Plante, and Asbjørnsen (2015) have applied similar Russian stimuli as the present experiment. However, they used single- and double marked words. The double-marked words in the experiment were gender marked by two suffixes that always co-occur (e.g., -tel + -ya). If a root word takes the masculine double marking –telya, it can also take the alternative marking –telyem. To get the feminine double-marked words the phoneme –k is combined with the feminine suffixes (i.e., –koj and –ku). This relationship between the suffixes can be considered an adjacent dependency, which is considered to be easier to learn than nonadjacent dependencies (Newport & Aslin, 2004). Several studies argue that multiple correlated cues assist the learner in successfully utilizing distributional information for category acquisition (Braine, 1992; Frigo & McDonald, 1998; McDonald & Plauché, 1995; Mintz, 2002; Newport & Aslin, 2004). Consistent with previous studies (e.g., Gerken, et al., 2005; Richardson, et al., 2006; Torkildsen, Dailey, Aguilar, Gómez, & Plante, 2013), Eidsvåg, et al. (2015) discovered that the double-marked words generated a larger effect of learning than single-marked words. Based on findings on the beneficial effect of variability for learning (Eidsvåg, et al., 2015; Plante et al., 2014; Richardson, et al., 2006; Torkildsen, et al., 2013), only double-marked words were included in the present experiment. This was in order to test distribution of exposure in an efficient manner and to reduce the impact of other sources of variance.

Following the publication of Williams (2005), the role of awareness in learning has become a debate within the field of second language acquisition. The study offered evidence
for the possibility to learn without awareness, in contrast to an earlier study by Leow (2000), reporting that learning without awareness did not appear to play an important role in either second or foreign language development. Noteworthy, the discrepancy between these two studies can be explained by methodological differences where the research designs measured unawareness at different stages of the learning process. Rebuschat, Hamrick, Riestenberg, Sachs, and Ziegler (2015)’s experiment sought to contribute to the current debate surrounding learning without awareness by triangulating measures of awareness. Their objective was to determine the advantages and disadvantages of multiple measures of awareness (concurrent verbal reports, retrospective verbal reports, and subjective measures of awareness). The study confirmed that learners are able to rapidly acquire novel form-meaning connections under incidental learning conditions and without the benefit of feedback (Rebuschat, et al., 2015). In terms of the different measurements of awareness, the inclusion of the think-aloud procedure occasionally revealed that awareness emerged earlier than the participants reported in the interviews, indicating that retrospective recall may be unreliable, but still useful in that they revealed partial rules that participants may have formed which could explain their performance (Rebuschat, et al., 2015). However, this could also be the cause of prompting participants to verbally describe rules or patterns at the end of the experiment. Additionally, the subjective measures of awareness allowed the detection of implicit and explicit knowledge (Rebuschat, et al., 2015).

Participants’ degree of awareness is considered essential for many researchers and seems to have an effect on learning, but the exact relationship between learning and awareness is more or less unknown (Frensch & Runger, 2003; Rebuschat, et al., 2015). While implicit learning studies usually contain measurements of awareness this is not typically featured in the statistical paradigm (Hamrick & Rebuschat, 2011). Subjective measurements of awareness can provide researchers with rich characterizations of the knowledge
participants acquire during experiments (Rebuschat & Williams, 2011). Since language comprehension and production are thought to be partly based on implicit knowledge, it seems important to determine if subjects in statistical learning research develop this type of knowledge. The lack of awareness is often assumed, but not empirically assessed (Aslin, Saffran, & Newport, 1999) leading to uncertainty whether statistical learning typically results in conscious or unconscious knowledge (Hamrick & Rebuschat, 2011). Basic measurements of awareness could be administrated since the present experiment was conducted on normally developed adults. By incorporating a self-rating questionnaire as a measurement of awareness, we could get an idea of the degree to which the participants engaged in hypothesis-testing or used other forms of explicit strategies during the present experiment.

The present experiment’s purpose was to investigate whether distribution of exposure might facilitate or hinder the learning of noun gender subcategories in an unfamiliar, natural language (Russian). Our main hypothesis was that a learning effect would occur, measured by acceptance of grammatical and ungrammatical items. The second hypothesis was that if learning did occur, we would see a difference in the distribution of exposure between the groups, where distributed condition would be most beneficial for learning based on the fact that an appreciable amount of research arguments in favor of distributed frequency of treatment scheduling. Finally, if learning occurred, we wanted to investigate if the participants were aware or able to explicit express what they had learned.

Method

Participants

Forty graduate students were included as voluntary participants based on a signed consent. They consisted of 20 men ($M = 22.75; SD = 2.07$) and 20 women ($M = 22.75; SD = 2.61$). None of them had any knowledge of Russian or other languages similar to Russian (e.g., Slavic languages). The participants included in the experiment were the ones that
responded first and met the inclusion criteria. However, participants who reported difficulties considering language learning in general, developmental or acquired neurological disorders, or had a known hearing loss were excluded from the experiment. In addition, the participants’ language background was collected.

The participants were primary Norwegian speakers and only four participants (two in the massed condition and two in the distributed condition) reported a different native language. However, all the participants reported that they understood and spoke Norwegian fluently. Eleven participants reported knowledge of one language in addition to their native language (four in the massed condition; seven in the distributed condition). The majority of the participants had knowledge about two or more languages (13 in the massed condition; 10 in the distributed condition) in addition to their native language. The remaining participants reported knowledge about three or more languages. The highest number of additional languages spoken was six (in the massed condition). Despite that the participants reported some differences in previous language knowledge, there was no significant difference between the numbers of languages the participants reported knowing between the two groups: massed condition: $M = 2.25, SD = 1.20$; distributed condition: $M = 1.75, SD = 0.63$.

**Experimental Design and Stimuli Material**

The present study had a between-within subject design where the participants were assigned to the two conditions (massed and distributed) by pseudo-randomization to ensure equal numbers of men and women in each condition. By dividing the participants into two conditions, we were able to investigate whether distribution of exposure influences initial learning of the Russian noun gender marking system. The auditory stimuli were delivered either divided in three cycles of 64 items per cycle separated by a test phase (distributed condition) or in three cycles consecutively of 64 items (in total 192 items) before the test phase (massed condition). Regardless of the differences in stimuli presentation (distributed vs.
massed), participants in both conditions were exposed to the same amount of stimuli (total of 192 items) and only differed in the presentation of the stimuli (i.e., all stimuli presented in one cycle before the test phase, or separated in three cycles with the test phase in between).

While none of the participants spoke Russian or another Slavic language, their language background might have affected their learning abilities when it came to the gender marking system of the Russian language. Similar to Norwegian, the Russian noun gender system contains three subcategories (masculine, feminine and neuter; Corbett, 1991), but in the present experiment only the subsets of masculine and feminine gender markings were included. It is common for Norwegians to be multilingual as it is generally known that Norwegian, Danish and Swedish are mutually understood and English is an obligatory second language for Norwegians. In addition, it is common to choose an elective third language during secondary school. The participants with knowledge of other languages may have an advantage of being more sensitive to the diversity in distributional cues that occur across languages in general.

**Familiarization stimuli.** The familiarization phase contained 32 different grammatically correct items with double-marked suffixes (64 unique items), listed in Table 1. Double-marked words indicate that two diminutive suffixes rather than one mark the gender. The gender markings were the stem ending and the case inflection (e.g., Rastochi - tel+ yem, underlined letters being the case inflection). There were two variations of masculine double markings (tel + ya and tel + yem) and two variations of feminine double markings (k + oj and k + u). Any root word that could take one of the two possible feminine endings (e.g., Blondin + koj) could also take the other feminine ending (e.g., Blondin + ku). Similar, the masculine root word ending –telya could use the alternative ending –telyem (e.g., Rastochi + -telya; Rastochi + -telyem). The two conditions consisted of an equal number of items (64 combinations of root words and gendered suffixes) that was delivered either as three cycles
consecutively of 64 items (in total 192 items) before the test phase (massed condition) or divided into three cycles of 64 items separated by a test phase (distributed condition). Each cycles had an individual new random order. Despite the differences in the stimuli presentation, the participants in both groups heard the same total amount during the experiment. If the participants were to perform well in the test phase, they had to detect the connection between the paired suffixes and the combination with the root words. In addition, the stimuli were presented in a randomized order. The intention of having a randomized order was to emphasize the way one naturally could come across words when learning an unfamiliar, natural language.

The auditory material was recorded by a male and a female who were native speakers of Russian. Previous studies indicate that multiple talkers promote learning compared to a single talker (e.g., Richtsmeier, Gerken, Goffman, & Hogan, 2009; Rost & McMurray, 2009). The sound files used in the present experiment were some of the same recordings of double-marked Russian words as in Eidsvåg, et al. (2015) with a few additional double-marked words. All the words had been edited to correspond to the actual length of the word (silent intervals were removed) and to produce relatively equal loudness across words. There was a 300-580 milliseconds inter-stimulus interval between each word.

| Table 1 here |

**Test phase.** In this phase, the participants in both groups heard 24 new individual items (listed in Table 2) and were asked to judge whether the items were grammatically correct or -incorrect Russian words. In total, the participants in both groups heard 72 items. The stimuli were delivered either as three cycles consecutively with 24 items (massed condition) or divided into three cycles of 24 items separated by the familiarization phases (distributed condition). Each cycles had an individual new random order. In the test phase
each root word had either a grammatical correct ending (e.g., Deja – telyem) or an ungrammatical word ending (e.g., Deja – telu). Half of the items were grammatically correct (i.e., grammatical items), while the other half violated the grammatical pattern (i.e., ungrammatical items) they had been exposed to during the familiarization phase. On the computer screen they were asked to click on the symbol of a “smiley face” or “frowny face” to mark a test item as correct or incorrect respectively, in order to continue to the next item. Hearing any other ending than the four suffixes from the familiarization phase should have lead the participants to reject the test item. A different female native speaker of Russian than heard in the familiarization phase recorded all these words.

Table 2 here

**Self-rating questionnaire.** After completing the experiment the participants filled out a self-rating questionnaire that was previously used by Eidsvåg, et al. (2015) (see Appendix A). In the questionnaire form we asked about the participants’ mother tongue and additional languages known. We also asked about whether the participants had any knowledge about Russian grammar. The information regarding the participants’ self-reported degree of awareness, rule searching, randomness and assertiveness during testing were quantified in order to see whether there was any correlation between degree of awareness and acceptance of grammatical versus ungrammatical items. The questionnaire form was also designed to go from general questions about awareness (i.e., to what extent did you feel that the answers you gave were random?) to more specific questions (i.e., to what extent will you assume that “-yem” followed “-tel”?). The answers were graded from 0 to 3 (0 = not aware/random to a very large extent/did not follow rules, 3 = aware/not random/followed rules). Information concerning the participants’ age was also collected.
Minor changes were done to the self-questionnaire form previously used in Eidsvåg, et al. (2015). We changed the wording of some of the questions to make them less ambiguous and easier to interpret, and we also added supplementary questions with increased specificity. One of the questions asked the participants to give six suffixes (three grammatical and three ungrammatical) a score from 0-100 by their thought of probability of the suffix occurred during the experiment. The mean for the grammatical correct items and the mean for the ungrammatical items were made into variables respectively.

Procedure

The research procedures fulfilled the criteria by the Norwegian Center for Research Data (Norsk senter for forskningsdata, NSD) and it was therefore no need to apply or get an approval according to Norwegian regulations. The experiment was conducted in a research lab at the University of Bergen. Participants individually went through the experimental tasks after signing the informed consent form. They were seated in front of a laptop, asked to put on headphones, and given instructions on how to adjust the volume. The researcher left the room, but remained at hand in the event of any technical issues occurring. During the experiment, the participants were given written instructions on the computer prior to each familiarization phase and test phase. The computer-based part of the experiment lasted for approximately 15 minutes, with a self-rating questionnaire given out afterwards. The experiment as a whole lasted for about 30 minutes.

The participants were given a monetary compensation of NOK 150 (approximately USD 18) for taking part in the experiment intended to cover any transportation costs etc. The money was given through the app “Vipps by DNB” or in cash after the completion of the experiment.
**Data processing and analyses.** The stimuli presentation and data collection were performed with E-prime 2.0 Professional (Schneider, Eschmann, & Zuccolotto, 2002). The data was later analyzed with the use of Statistica version 13.

Acceptance rate (i.e. number of accepted grammatical versus ungrammatical items) and response time were identified as dependent variables. In experiments of this nature, acceptance rate is typically used where learning is operationalized as significantly higher acceptance of grammatical items compared to ungrammatical items.

**Results**

To address whether distribution of exposure had an impact on the participants’ ability to learn Russian noun gender system, a two (group: massed vs. distributed) x two (grammaticality: acceptance of grammatical vs. acceptance of ungrammatical) x three (cycles: 1, 2, 3) mixed analysis of variance (ANOVA) was conducted. The group condition (massed vs. distributed) was treated as a between-group factor and grammaticality and cycle were treated as repeated measurements within subjects. The acceptance rate data for each item type and condition are presented in Table 3 and visualized in Figure 1.

---

A statistically significant main effect for grammaticality occurred, where participants in both conditions primarily seemed to accept more grammatical items than ungrammatical items, $F(1,37) = 33.74, p = .00, \eta_p^2 = .47$. A significant main effect for cycle have also been found, $F(2,74) = 12.79, p = .00, \eta_p^2 = .25$. In both groups, the grammatical and ungrammatical items are gradually accepted more frequently after each cycle. A post hoc Tukey HSD test indicated more acceptance of grammatical and ungrammatical items after the
first cycle where the results of cycle 2 ($p = .00$) and cycle 3 ($p = .00$) showed that a learning tendency manifested after cycle 1. No other sources of variance yielded significant effects.

A 2 x 2 x 3 mixed ANOVA was conducted to see if there was any group effect on the participants’ response time within the three cycles. The response time data are presented in Table 4 and visualized in Figure 2. A statistically significant main effect for cycle was found $F_{(2, 74)} = 10.90, p = .00, \eta^2_p = .23$. The post-hoc Tukey HSD indicated that cycle 1 significant differed from cycle 2 and cycle 3 ($p = .00$). A significant effect for cycle and group condition was found, $F_{(2, 74)} = 4.99, p = .00, \eta^2_p = .12$. The massed condition had a significant difference between cycle 1 and cycle 2 as well as cycle 1 and cycle 3. Even though the participants in the distributed condition responded faster in cycle 3 compared to the massed condition, they did not accept nor reject more grammatical items. No other sources of variance yielded significant effects.

In order to investigate whether the participants’ response time had an impact on their performance, a correlational analysis was conducted (presented in Table 5). A positive correlation was revealed between the response time to ungrammatical items in cycle 1 and the acceptance of ungrammatical items in Cycle 2, $r = .36, p < .05$. Additionally, we found a correlation between the acceptance of grammatical items in the first cycle, and the response time of grammatical items in cycle 2, $r = .31, p < .05$, and response time for ungrammatical items in cycle 3, $r = .31, p < .05$. 
An independent samples T-test was conducted to compare the responses from the questionnaire for the two groups (massed and distributed). The t-test is presented in Table 6. While none of the scores revealed any significant differences, the questions about the participants’ probability for accepting grammatical and ungrammatical items were of interest. Participants in both groups were more inclined to accepted grammatical items (massed: $M = 72.79$, $SD = 15.05$, ns; distributed: $M = 74.00$, $SD = 15.71$, ns) than ungrammatical items (massed: $M = 44.30$, $SD = 12.56$, ns; distributed: $M = 38.67$, $SD = 16.65$, ns).

Correlations between participants’ performance and questionnaire responses are presented in Table 7. The degree to which participants felt their answers were random correlated negatively with acceptance of ungrammatical items in cycle 2 ($r = -.38$, $p < .05$) and in cycle 3 ($r = -.32$, $p < .05$). The degree to which the participants felt there was any patterns they could follow while solving the task correlated negatively with acceptance of ungrammatical items in cycle 3 ($r = -.32$, $p < .05$). The more specific question on the probability of ungrammatical words correlated positively with acceptance of grammatical items in cycle 1 ($r = .39$, $p < .05$). There were no significant correlations between the degree to which participants reported searching for regularities or reported probability of grammatical words and response accuracy. A between-within ANOVA was conducted to test for an interaction effect between the probability variables and the group conditions. The results did not yield any significant effects.

None of the participants could explicitly explain the rules that governed their response to the gender suffixes. However, more than 50% of the 40 participants stated that their attention had been drawn towards the suffixes. Six of the participants (four from the massed group and two from the distributed group) explained further by identifying patterns they had
recognized. One participant in the massed group reported recognizing suffixes ending with –lem, -kai, -o, -a, -koj. One other participant from the massed group reported scoring suffixes like –ov, -o, -em with a smiley face (i.e., grammatical correct) and words ending with –a and –i with a frowny face (i.e., grammatical incorrect). One of the participants from the distributed group believed suffixes like –ku, -ai, -yem, -tel, were correct. Another participants from the distributed group believed words ending with –a, did not fit in.

**Discussion**

The main objective of the present experiment was to investigate whether distribution of exposure would facilitate the learning of noun gender subcategories in an unfamiliar, natural language. The findings reveal that the participants in both conditions primarily seem to accept more grammatical items than ungrammatical items. This could indicate that they were able to learn the grammatical pattern and that our first hypothesis of there being an overall learning effect was supported. The abovementioned results also indicated that there was no significant difference in the acceptance of grammatical and ungrammatical items on a group level (i.e., between the massed condition and the distributed condition). Therefore, the second hypothesis that there would be a difference in learning between the massed and distributed condition was not supported. Literature on massed and distributed learning generally indicates that learning is enhanced by learning sessions distributed over a longer time period (e.g., Janiszewski, et al., 2003). For example, last minute studying before an exam is not as efficient as studying steadily over a longer time frame. Contrariwise, our findings are in accordance with other previous studies demonstrating that distribution of exposure indicates no difference between massed and distributed group performance (Meyers-Denman & Plante, 2016; Proctor-Williams & Fey, 2007; Ukrainetz, et al., 2009). A large body of experimental evidence actually suggests that both the distributed condition and the massed condition facilitate language learning (e.g., Ambridge, Theakston, Lieven, & Tomasello,
DISTRIBUTION OF EXPOSURE AND SUBCATEGORY LEARNING

2006; Barratt, et al., 1992; Riches, et al., 2005; Vlach, et al., 2008). While the participants in the present experiment received the same number of doses and the same number of stimuli over an equal period of time, the similar effect of distribution of exposure may be attributed to the manipulation occurred within 15 minutes for both conditions. Other treatment studies have investigated the intensity over a course of days or weeks (e.g., Barratt, et al., 1992; Smith-Lock, et al., 2013) with results indicating one condition giving greater learning improvements. A study over a longer period of time might have lead to different results leaning towards either a better learning effect for the massed condition or the distributed condition. Arciuli and Torkildsen (2012) pointed out that longitudinal data is needed to advance the understanding of the role that statistical learning might play in the process of language acquisition, both in the context of normal development and developmental disorders. Conversely, by having a study that examines the dynamics of a phenomenon over time a possible problem could be attrition, i.e., the loss of participants over time (Polit & Beck, 2012). Attrition becomes problematic if the participants who drop out differ in important ways from those who continue to participate. This could result in potential biases and difficulty with generalizing to the original population (Polit & Beck, 2012). Additionally, the studies mentioned about massed and spaced learning has generally focused on treatment (Barratt, et al., 1992; Meyers-Denman & Plante, 2016; Smith-Lock, et al., 2013), have used children (Barratt, et al., 1992; Meyers-Denman & Plante, 2016; Proctor-Williams & Fey, 2007; Riches, et al., 2005; Smith-Lock, et al., 2013; Vlach, et al., 2008), or have shown the difference between children with specific language impairments (Meyers-Denman & Plante, 2016; Proctor & Vu, 2003; Riches, et al., 2005; Smith-Lock, et al., 2013), not on incidental learning in adults. If the present experiment would have incorporated any of the abovementioned factors the results might have differed. On a group level there was a difference between the conditions, where the massed condition showed a higher learning
effect in the first cycle. A similar tendency appeared in the study by Eidsvåg, et al. (2015), where learning of grammatical items occurred in the first cycle, but declined in the second cycle. However, the participants accepted more grammatical items in the third cycle, but not at the same level as in the first cycle (Eidsvåg, et al., 2015). A speculation is that this may be explained by the exposure to ungrammatical items during the testing leading to more explicit strategies. Another explanation might suggest that the difference in distribution of stimuli affects the efficiency of a statistical learning mechanism and may predict how the knowledge retains over time. This development could be explained with how the nature of the task and context could trigger in different learning strategies (Riding & Rayner, 2013). Additionally, other studies have suggested that when facing different types of learning problems learners might use dissimilar strategies (Marcus, 2000; Peña, Bonatti, Nespor, & Mehler, 2002).

The divergence of response time between the two conditions in the first cycle and differing response time in cycle 2 and cycle 3 could indicate that they had different learning strategies. The correlations could indicate that the participants response time in the first cycle predicted their acceptance of grammatical and ungrammatical items in the later cycles. While the massed condition had the fastest response time in cycle 2, the distributed condition’s response time decreased steadily from first to the last cycle. Another alternative for the participants’ faster response time could be that the participants started to give up and then made faster decisions towards the end of the experiment. In test situations, individuals may attempt to maximize performance on the accuracy of their response and response time (Proctor & Vu, 2003). There are however situations were an individual increases his or her response time at the cost of reducing the accuracy, while in another situation he or she will need to slow down the response time in order to have a better accuracy level (Zimmerman, 2011). If the participants in the present study were aware of their response time being measured, it could mean they would try to maximize either their response or their response
time. In the present experiment the participants were, as far as we know, not aware of their response time being measured. An interesting aspect for further research would be to investigate how the participants’ response time would endure if the experiment had a fourth cycle. Their response time in the previous cycles might be an indication of what the participants would respond in the cycles after. This is only a speculation and the conclusion is that there was no speed accuracy tradeoff, neither at a group level nor in the cycle.

Similar to previous studies (Eidsvåg, et al., 2015; Gerken, et al., 2005; Torkildsen, et al., 2013), the double-marked words showed a learning effect when it comes to learning a subcategory system measured by the participants acceptance of more grammatical items than ungrammatical items. The Russian double-marked words used in the present experiment had an adjacent dependency, which is seen to be easier to learn than non-adjacent dependencies (Newport & Aslin, 2004). Several studies have shown that multiple cues could be the reason for a better learning effect occurring (Braine, 1992; Frigo & McDonald, 1998; McDonald & Plauché, 1995; Mintz, 2002; Newport & Aslin, 2004). In the present experiment, the results could indicate that a learning tendency manifested after the first cycle, since the learning of grammatical items was maintained throughout cycle 2 and cycle 3. Acceptance of grammatical items increased during the whole of the experiment and it would appear that the grammatical items were a more probable response than ungrammatical items. However, this was also the case for the ungrammatical items. The participants’ response could be caused by a number of factors, since any information in the experiment setting could generate a certain response. The participants might have been inclined to agree to all the test items. Firstly, this could be because the participants want to be agreeable and not disappoint the researcher (Polit & Beck, 2012). Secondly, it could be caused by biases in memory where are more inclined to remember information in agreement than is a contradiction (Cronbach, 1942). Lastly, this
increased acceptance of grammatical and ungrammatical items in could also be a sign of boredom and wanting to end the experiment.

Another possible explanation for the learning effect could be that the participants’ language background influenced the performance. We controlled for the effect of prior language knowledge and language transfer by excluding participants with knowledge of Russian or other Slavic languages. Moreover, the participants’ previous language knowledge did not correlate significantly with their responses or response time. However, there are some similarities between the grammars of Norwegian and Russian (e.g., both languages having three subcategories for gender marking: feminine, masculine, neuter) that should be noted. The similarities could make the participants more perceptive of ways distributional cues are in other languages. Instead of measuring the statistical learning mechanism we are after, we could be tapping into a general language knowledge or a sensitivity to certain aspects of the linguistic input based on previous linguistic knowledge.

Our final hypothesis was to investigate if the participants were able to explicit express what they had learned if a certain form of learning occurred. With the use of the self-questionnaire form we wanted to determine the participants’ level of awareness regarding the rules for subcategorical gender marking. This was administered after the completion of the experiment and could therefore not be used to compare their awareness throughout the cycles. Comparing the participants’ degree of awareness revealed that there was no difference on a group level (massed vs. distributed). The way in which a question is asked may trigger different response. Therefore, by going from general to more specific questions in the questionnaire, we wanted to investigate where on the continuum of awareness the participants became conscious of regularities and patterns in the Russian noun gender system. It was probably difficult for the participants to express what they had gathered of knowledge explicitly, but they reported that they recognized patterns and paid attention to the word
endings. This could imply that some degree of explicit hypothesis testing occurred during the experiment. If the participant’s hypothesis was correct, looking for patterns and regularities in the word endings might have facilitated learning, however, there should be sufficient stimuli material to enhance the regularities (Cleeremans, Destrebecqz, & Boyer, 1998). When forced to give each of the endings presented in the questionnaire a score between zero and one hundred of whether the suffixes presented had occurred in the experiment or not, participants in both conditions seemed more inclined to accept the grammatical items heard during the experiment and reject the ungrammatical items presented in the test phase. In addition, the participants from both conditions reported that they were fairly certain that they heard –koj and -telyem. These results can indicate that some degree of explicit awareness for the Russian gender marking system occurred and that there was a tendency for the participants to discover some patterns under incidental learning conditions without the benefit of feedback. The non-significant results could have been influenced by the participants testing out different hypothesis before arriving at the correct hypothesis they finally report in the questionnaire (e.g., noticing the suffixes as a pattern to pay attention to). However, retrospective recall can be unreliable, even if it is useful to reveal partial rules that participants formed that could explain their performance (Rebuschat, et al., 2015). Perhaps if a think-loud procedure had been included, the measurement of awareness might have been better to investigate when, or if, the participants began testing hypotheses or used explicit strategies during the experiment. The questionnaire did allow us to investigate if the participants had any awareness.

However, when interpreting the participants’ self-reported awareness from the more specific question regarding the probability of suffixes occurring in the experiment, we had to be caution when making assumptions since there are many factors that could generate a response in an experiment setting. For instance, one explanation for the relative difference
could be caused by the presentation of the input, since the participants were exposed to auditory input during the experiment while the questionnaire was in written.

**Conclusion**

To summarize, the results indicate that differences in distribution did not have a significant effect and there was no difference between the massed condition and the distributed condition. The results extend the evidence base by reporting the effect of distribution of learning noun gender subcategories in an unfamiliar, natural language. Moreover, the participants in both conditions primarily seemed to accept more grammatical items than ungrammatical items, which could indicate a learning effect. While a speed-accuracy tradeoff was not found, neither on a group level nor across the cycles, the interaction between acceptance rate and response time could indicate that the participants used different learning strategies throughout the experiment. The findings from the questionnaire revealed that the massed and distributed conditions reported similar degree of awareness and did not differ on a group level during the experiment. In relation to the more specific questions in the questionnaire, there was a tendency that the participants could have had an effect of learning showed by giving grammatical items a higher probability than ungrammatical items.

Results from the presented study may enlighten researchers interested in the appearance of distribution of exposure in statistical learning experiments, and extend the research by promoting the investigation of distribution of exposure in practice and across different populations. To improve the quality of language treatment designs it is necessary to reduce the differences between theory and practice, and it is crucial that speech-language pathologists understand the principles of learning better (Alt, Meyers, & Ancharski, 2012). The main finding of our study contributes to the debate around distribution of exposure in language acquisition, and can therefore be part of a point of discussion for both researchers
and clinicians. In addition, the present experiment sheds light on the intricate relationship between statistical learning and awareness.
References


Norsk senter for forskningsdata, N. *Notification form for research- and student projects processing personal data*. Retrieved from

http://www.nsd.uib.no/personvernombud/meld_prosjekt/meldeskjema?eng


doi:10.1044/2014_AJSLP-13-0038


Table 1

*Stimuli Set for the Familiarization Phase*

<table>
<thead>
<tr>
<th></th>
<th>Masculine words</th>
<th>Feminine words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female voice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dviga -telya/telyem*</td>
<td></td>
<td>Blondin -koj/ku*</td>
</tr>
<tr>
<td>Khrani –telya/telyem</td>
<td></td>
<td>Kodoerov –koj/ku</td>
</tr>
<tr>
<td>Muchi-telya/telyem*</td>
<td></td>
<td>Progul –koj/ku</td>
</tr>
<tr>
<td>Osnova-telya/telyem*</td>
<td></td>
<td>Rozoch –koj/ku</td>
</tr>
<tr>
<td>Sluzhi –telya/telyem</td>
<td></td>
<td>Rubash –koj/ku*</td>
</tr>
<tr>
<td>Smotri –telya/telyem</td>
<td></td>
<td>Skovorod-koj/ku*</td>
</tr>
<tr>
<td>Sozda –telya/telyem</td>
<td></td>
<td>Vetrov –koj/ku</td>
</tr>
<tr>
<td>Uchi –telya/telyem</td>
<td></td>
<td>Vystava –koj/ku</td>
</tr>
<tr>
<td><strong>Male voice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blagode - telya/telyem*</td>
<td></td>
<td>Babush –koj/ku</td>
</tr>
<tr>
<td>Dushi -telya/telyem*</td>
<td></td>
<td>Devoch –koj/ku</td>
</tr>
<tr>
<td>Grabi –telya/telyem</td>
<td></td>
<td>Juboch –koj/ku</td>
</tr>
<tr>
<td>Pisa –telya/telyem</td>
<td></td>
<td>Karmel-koj/ku*</td>
</tr>
<tr>
<td>Potrebi –telya/telyem</td>
<td></td>
<td>Khlopush –koj/ku</td>
</tr>
<tr>
<td>Rastochi –telya/telyem</td>
<td></td>
<td>Makush -koj/ku*</td>
</tr>
<tr>
<td>Razaruchi –telya/telyem</td>
<td></td>
<td>Podush –koj/ku</td>
</tr>
<tr>
<td>Vodi -telya/telyem</td>
<td></td>
<td>Petrush-koj/ku</td>
</tr>
</tbody>
</table>

*Note. The words were recorded equally by two native speakers of Russian, a male speaker and a female speaker. * = indicates stimulus previously used in Eidsvåg et al. (2015).*
Table 2

*Stimuli Set for the Test Phase*

<table>
<thead>
<tr>
<th></th>
<th>Masculine words</th>
<th>Feminine words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grammatical/correct</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>words</td>
<td>Deja –telyem</td>
<td>Brjunet –koj*</td>
</tr>
<tr>
<td></td>
<td>Osvezhi –telyem*</td>
<td>Obert –koj</td>
</tr>
<tr>
<td></td>
<td>Pokupa –telyem*</td>
<td>Tarel –koj</td>
</tr>
<tr>
<td></td>
<td>Rodi –telya</td>
<td>Devush –ku*</td>
</tr>
<tr>
<td></td>
<td>Stroi –telya</td>
<td>Maka –ku*</td>
</tr>
<tr>
<td></td>
<td>Ljubi –telya*</td>
<td>Verev –ku</td>
</tr>
<tr>
<td><strong>Ungrammatical/incorrect</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>words</td>
<td>Deja –telu</td>
<td>Brjunet -kya</td>
</tr>
<tr>
<td></td>
<td>Osvezhi –telyem*</td>
<td>Obert –kya</td>
</tr>
<tr>
<td></td>
<td>Pokupa –telyem*</td>
<td>Tarel –kya</td>
</tr>
<tr>
<td></td>
<td>Rodi –teloj</td>
<td>Devush –kyem*</td>
</tr>
<tr>
<td></td>
<td>Stroi –teloj</td>
<td>Maka –kyem*</td>
</tr>
<tr>
<td></td>
<td>Ljubi –teloj*</td>
<td>Verev –kyem</td>
</tr>
</tbody>
</table>

*Note. The words were recorded by a different female Russian speaker than in the familiarization phase. * = indicates stimulus previously used in Eidsvåg et al. (2015).*
Table 3

Descriptive Statistics for Acceptance Rate Data (Accepted Grammatical vs. Ungrammatical Items) for the Massed and Distributed Condition in the Three Cycles

<table>
<thead>
<tr>
<th></th>
<th>Massed Condition</th>
<th>Distributed Condition</th>
<th>t-value</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle 1: Grammatical</td>
<td>7.75 (2.12)</td>
<td>7.25 (1.40)</td>
<td>0.88</td>
<td>38</td>
<td>0.39</td>
</tr>
<tr>
<td>Cycle 1: Ungrammatical</td>
<td>5.15 (1.66)</td>
<td>5.50 (1.60)</td>
<td>-0.68</td>
<td>38</td>
<td>0.50</td>
</tr>
<tr>
<td>Cycle 2: Grammatical</td>
<td>8.15 (2.23)</td>
<td>8.15 (2.32)</td>
<td>0.00</td>
<td>38</td>
<td>1.00</td>
</tr>
<tr>
<td>Cycle 2: Ungrammatical</td>
<td>6.00 (2.08)</td>
<td>6.35 (1.95)</td>
<td>-0.54</td>
<td>37</td>
<td>0.59</td>
</tr>
<tr>
<td>Cycle 3: Grammatical</td>
<td>8.00 (2.49)</td>
<td>8.65 (1.72)</td>
<td>-0.96</td>
<td>38</td>
<td>0.34</td>
</tr>
<tr>
<td>Cycle 3: Ungrammatical</td>
<td>6.35 (2.51)</td>
<td>6.90 (2.17)</td>
<td>-0.74</td>
<td>38</td>
<td>0.46</td>
</tr>
</tbody>
</table>
Table 4

*Descriptive Statistics of Response Time under the Massed Condition and the Distributed Condition.*

<table>
<thead>
<tr>
<th></th>
<th>Massed Condition</th>
<th>Distributed Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>( M )</td>
</tr>
<tr>
<td>Cycle 1: GA</td>
<td>885.29 (536.80)</td>
<td>1243.04 (574.75)</td>
</tr>
<tr>
<td>Cycle 1: UGA</td>
<td>1115.61 (673.50)</td>
<td>1532.41 (862.13)</td>
</tr>
<tr>
<td>Cycle 2: GA</td>
<td>729.42 (461.76)</td>
<td>832.83 (518.23)</td>
</tr>
<tr>
<td>Cycle 2: UGA</td>
<td>690.83 (381.40)</td>
<td>1070.20 (754.95)</td>
</tr>
<tr>
<td>Cycle 3: GA</td>
<td>941.63 (644.46)</td>
<td>806.06 (570.24)</td>
</tr>
<tr>
<td>Cycle 3: UGA</td>
<td>899.19 (602.65)</td>
<td>792.45 (463.48)</td>
</tr>
</tbody>
</table>

*Note.* GA = Acceptance of grammatical items; UGA = Acceptance of ungrammatical items
Table 5

*Correlations Between Acceptance Rate and Response Time*

<table>
<thead>
<tr>
<th></th>
<th>Cycle 1: RT for GA</th>
<th>Cycle 1: RT for UGA</th>
<th>Cycle 2: RT for GA</th>
<th>Cycle 2: RT for UGA</th>
<th>Cycle 3: RT for GA</th>
<th>Cycle 3: RT for UGA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle 1: GA</td>
<td>0.06</td>
<td>0.03</td>
<td>0.32</td>
<td>0.10</td>
<td>0.08</td>
<td>0.32</td>
</tr>
<tr>
<td>Cycle 1: UGA</td>
<td>0.03</td>
<td>0.20</td>
<td>0.02</td>
<td>0.04</td>
<td>0.25</td>
<td>-0.08</td>
</tr>
<tr>
<td>Cycle 2: GA</td>
<td>0.16</td>
<td>0.10</td>
<td>0.18</td>
<td>-0.12</td>
<td>-0.12</td>
<td>0.29</td>
</tr>
<tr>
<td>Cycle 2: UGA</td>
<td>0.27</td>
<td>0.36</td>
<td>-0.05</td>
<td>0.17</td>
<td>0.07</td>
<td>-0.08</td>
</tr>
<tr>
<td>Cycle 3: GA</td>
<td>0.27</td>
<td>0.18</td>
<td>0.11</td>
<td>0.14</td>
<td>-0.06</td>
<td>0.26</td>
</tr>
<tr>
<td>Cycle 3: UGA</td>
<td>0.21</td>
<td>0.19</td>
<td>0.06</td>
<td>-0.06</td>
<td>0.04</td>
<td>0.02</td>
</tr>
</tbody>
</table>

*Note.* Marked correlations are significant at p < .05000. RT = response time; GA = acceptance of grammatical items; UGA = acceptance of ungrammatical items.
Table 6

*T-test for the Questionnaire*

<table>
<thead>
<tr>
<th>Question</th>
<th>Massed Condition</th>
<th>Distributed Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>1. Degree of randomness</td>
<td>1.47</td>
<td>0.70</td>
</tr>
<tr>
<td>2. Degree of awareness</td>
<td>1.63</td>
<td>0.76</td>
</tr>
<tr>
<td>3. Awareness of regularities</td>
<td>1.26</td>
<td>0.87</td>
</tr>
<tr>
<td>4. Correct patterns?</td>
<td>1.00</td>
<td>0.77</td>
</tr>
<tr>
<td>5. Mean score for probability of accepting</td>
<td>72.79</td>
<td>15.05</td>
</tr>
<tr>
<td>grammatical suffixes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Mean score for probability of accepting</td>
<td>44.30</td>
<td>12.56</td>
</tr>
<tr>
<td>ungrammatical suffixes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The questions were as follows: 1. How random were your responses? 2. Did you search for regularities? 3. Did you find any regularities? 4. Did you think that the patterns you might have found are correct? The answers were graded from 0 to 3 (0 = aware/not random/followed rules, 3 = not aware/random to a very large extent/did not follow rules). Questions 5 and 6 were graded from 0-100 (0 = not sure whether the suffix occurred in the experiment, 100 = sure that the suffix occurred in the experiment).
Table 7

Correlation for Acceptance Rate and the Questionnaire Answers

<table>
<thead>
<tr>
<th></th>
<th>Cycle 1</th>
<th>Cycle 2</th>
<th>Cycle 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GA  UGA</td>
<td>GA  UGA</td>
<td>GA  UGA</td>
</tr>
<tr>
<td>Age</td>
<td>0.17 -0.23</td>
<td>-0.05 -0.11</td>
<td>0.13 -0.13</td>
</tr>
<tr>
<td>No. of languages</td>
<td>0.10 -0.27</td>
<td>-0.09 -0.13</td>
<td>-0.02 -0.29</td>
</tr>
<tr>
<td>1. Degree of randomness</td>
<td>-0.12 -0.16</td>
<td>-0.15 -0.38</td>
<td>-0.29 -0.32</td>
</tr>
<tr>
<td>2. Degree of awareness</td>
<td>0.04 -0.23</td>
<td>-0.12 -0.17</td>
<td>0.05 -0.11</td>
</tr>
<tr>
<td>3. Awareness of regularities</td>
<td>-0.06 -0.31</td>
<td>0.05 -0.12</td>
<td>0.09 -0.32</td>
</tr>
<tr>
<td>4. Correct patterns?</td>
<td>-0.13 0.09</td>
<td>0.01 -0.00</td>
<td>0.16 -0.05</td>
</tr>
<tr>
<td>Probability of accepting grammatical suffixes</td>
<td>0.22 0.00</td>
<td>0.28 0.03</td>
<td>0.29 0.24</td>
</tr>
<tr>
<td>Probability of accepting ungrammatical suffixes</td>
<td>0.21 0.40</td>
<td>-0.17 0.26</td>
<td>0.03 0.28</td>
</tr>
</tbody>
</table>

Note. None of the correlations were significant at p < .05000. GA = acceptance of grammatical items, UGA = acceptance of ungrammatical items.
Figure 1. Acceptance rate for grammatical and ungrammatical items for the conditions across cycles. The small bars indicate standard error.
Figure 2. Response time for the massed and distributed condition across the three cycles. The small bars indicate standard error.
Språkkunnskaper (ring rundt det svaralternativet som passer deg)

(Language knowledge)

1. Hva er ditt morsmål?

(What is your native tongue?)

NORSK ANNET: __________

(Norwegian) (Other)

2. Hvilke språk kan du i tillegg til morsmålet? Hvor godt kan du disse?

(Which languages are you familiar with besides your first language? How well do you know these?)

3. Har du kjennskap til russisk grammatikk?

(Do you have any knowledge of Russian grammar?)

Ikke i det hele tatt Litt I stor grad I svært stor grad

(not at all) (some) (to a large extent) (to a very large extent)
De databaserte oppgavene (ring rundt det svaralternativet som passer deg)

(Computer-based tasks)

4. I hvilken grad føler du at svarene du ga var tilfeldige?

(To what extent do you feel the answers you gave were random?)

Ikke i det hele tatt   Litt   I stor grad   I svært stor grad
(not at all)   (some)   (to a large extent)   (to a very large extent)

5. Hvor bevisst var du på å lete etter regelmessigheter i det du hørte?

(How aware were you searching for regularities in what you heard?)

Ikke bevisst   Litt bevisst   Bevisst i stor grad   Bevisst i svært stor grad
(not aware) (aware to some extent) (aware to a large extent) (aware to a very large extent)

6. Opplevde du at det var noen mønster som du kunne følge i oppgaveløsningen?

(Did you feel that there was any pattern that you could follow in the performance of tasks?)

Ikke i det hele tatt   Litt   I stor grad   I svært stor grad
(not at all)   (some)   (to a large extent)   (to a very large extent)

Evt. hvilke? _______________________________________

(if so, which?)

7. I hvilken grad mener du disse mønstrene du eventuelt fulgte var korrekte?

(To what extent do you think these patterns you possibly followed were correct?)

Ikke i det hele tatt   Litt   I stor grad   I svært stor grad
(not at all)   (some)   (to a large extent)   (to a very large extent)
### Konkrete spørsmål

*(Specific questions)*

8. **I hvilken grad vil du anta at ”-yem” fulgte etter ”-tel”?**

   *(To what extent would you assume that ”-yem” followed ”-tel”)?*

<table>
<thead>
<tr>
<th>Ikke i det hele tatt</th>
<th>Litt</th>
<th>I stor grad</th>
<th>I svært stor grad</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>(not at all)</em></td>
<td><em>(some)</em></td>
<td><em>(to a large extent)</em></td>
<td><em>(to a very large extent)</em></td>
</tr>
</tbody>
</table>

9. **Vil du si det er noen sammenheng mellom ”-koj” og ”-ku”?**

   *(Would you say there is any correlation between ”-koj” and ”-ku”)?*

10. **Oppfattet du noen av endingene som mer passende?** *(ring rundt den/de)*

   *(Do you preceive any of the endings as more appropriate?)*

   -teloj  -ku  -telya  -kyem
11. Gi hver av disse endelsene en skår mellom null og hundre etter hvor høy
sannsynligheten er for at den forekom i listene du hørte 0 = sikker på at den ikke
forekom; 100 = helt sikker på at den forekom)

(Give each of these endings a score between zero and one hundred by how high the
probability is that it occurred in the lists you heard. 0 = sure it did not occur; 100 =
sure that it occurred)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-koj</td>
<td>-bovu</td>
<td></td>
</tr>
<tr>
<td>-kya</td>
<td>-telva</td>
<td></td>
</tr>
<tr>
<td>-telyem</td>
<td>-kovju</td>
<td></td>
</tr>
<tr>
<td>-telu</td>
<td>-selyem</td>
<td></td>
</tr>
</tbody>
</table>
Appendix B

Samtykke til deltagelse i studien

Jeg har mottatt informasjon om studien, og er villig til å delta

(Signert av prosjektdeltaker, dato)

Alder: ____

Kjønn: mann kvinne

Studieretning/yrke: ___________________
Appendix C

Instructions given to the participants on the computer screen before each phase.

*Welcome to this experiment!*

We are investigating what is called “implicit language learning”. Implicit language learning is when we unintentionally learn something about a language just by listening to it.

*Press SPACE to continue*

*You will now hear a series of words.*

You have to listen carefully, because afterwards we will check if you have learned them.

*Press SPACE when you are ready.*)
Instructions given to participants in the Distributed Condition

Nå skal du få høre noen av ordene en gang til.

Trykk på MELLOMROM når du er klar

(You will now hear some more words. Press SPACE when you are ready.)

Instructions before the Test Phase

Nå skal du få høre noen nye ord.

Du skal vurdere om ordene du nå hører er korrekte ord eller ikke.
Klikk smilefjes for riktige svar, og surt fjes for feile svar.

Trykk på MELLOMROM når du er klar

(You should decide whether the words you hear are correct words or not.
Press “smiley face” for correct words and “frowney face” for incorrect words.)

(Press SPACE when you are ready)
Test Phase

DET VAR DET HELE!
Takk for at du ble med!

(That's all!
Thank you for participating!)
Appendix D

Forespørsel om deltakelse i forskningsprosjekt

"Implisitt språklæring"

Bakgrunn og formål
Du har blitt invitert til å delta i et eksperiment som skal undersøke det som kalles implisitt språklæring. Implisitt språklæring handler om at vi ubevisst lærer noe om et språk vi ikke kjenner bare ved å lytte til det. Forsøket er en del av et mastergradsprosjekt ved Universitetet i Bergen.

Hva innebærer deltakelse i studien?

I forkant av studien blir det gitt informasjon om studiens eksklusjonskriterier. Personer med nevrologiske vansker, språkvansker eller hørselsvansker blir ikke inkludert i studien, dette gjelder også personer med kjennskap til russisk.

Hva skjer med informasjonen om deg?
All informasjon som samles inn vil bli behandlet konfidensielt og resultatene fra forsøket vil ikke kunne identifiseres i en eventuell publikasjon. Innsamlede data vil bli merket med en anonym kode og vil kun være tilgjengelige for de involverte mastergradsstudentene og veileder.


Frivillig deltakelse
Deltakelse i denne studien er helt frivillig og du kan på hvilket som helst tidspunkt trekke deg fra studien, uten å oppgi noen grunn. Dette vil ikke få noen negative konsekvenser for deg. Dersom du trekker deg, vil alle opplysninger om deg bli slettet.

Kontaktinformasjon
Ansvarlig for forsøket:
Professor Arve Asbjørnsen (arve.asbjornsen@psybp.uib.no).
Sunniva Sørhus Eidsvåg (ss099@ansatt.uib.no)

Mastergradsstudenter:
Kristina Tveit Osebakken (kristina.osebakken@student.uib.no)
Iselin Partee (iselin.partee@student.uib.no)