Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Tammy Lu April 10, 2024

Increased low-certainty modal verb use in English relative to Dutch may decrease subjective representations of the probability of future events

By

Tammy Lu

Dr. Phillip Wolff
Adviser

Linguistics

Dr. Phillip Wolff
Adviser

Dr. Marjorie Pak
Committee Member

Dr. Philip Kragel
Committee Member

2024

Increased low-certainty modal verb use in English relative to Dutch may decrease subjective representations of the probability of future events

By

Tammy Lu

Dr. Phillip Wolff
Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Linguistics

2024

Abstract

Increased low-certainty modal verb use in English relative to Dutch may decrease subjective representations of the probability of future events

By Tammy Lu

This study examines future time reference (FTR) in Dutch and English speakers and its effects on subjective representations of the probability of future events. FTR refers to any linguistic utterance that refers to a future event (Chen, 2013; Robertson, 2022). Dutch speakers, whose language doesn't oblige the usage of FTR to speak about future events, value future events more than English speakers, whose language obliges the usage of FTR. The hypothesized reason is that English speakers use more low-certainty verbs when referencing future events, causing them to have less subjective feelings about the probability of future events and value them less. A FTR elicitation task and a subjective certainty rating task were used to test the hypotheses. Participants were asked to conjugate main verbs according to the context of the situation and information about the likelihood of the situation happening. They were then asked to look at a series of images with differing proportions of red and blue dots. After, participants were asked how likely a random dot selected from those images would be red. The results offered inconclusive support for the hypotheses. Dutch and English speakers differed in their usage of low-certainty modal verbs with English speakers using more low-certainty modal verbs. The mediation analysis results showed an effect that went in the predicted direction but fell short of significance. It is suggested the study may be underpowered. This study provides an initial glimpse into how talk about future events might impact people's judgments the probability of their occurrence.

Increased low-certainty modal verb use in English relative to Dutch may decrease subjective representations of the probability of future events

By

Tammy Lu

Phillip Wolff
Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts Honors

Linguistics

2024

Acknowledgements

My thanks to my Phillip Wolff, my adviser, and Cole Robertson, the post-doctoral fellow who guided through this project, for helping me explore linguistic relativity, decision-making behavior, and non-linguistic cognition. I am grateful for your reliable advice and allowing me to gain a deeper insight on research experience.

Table of Contents

Introduction

The division of time into the past, present, and future is a commonly observed set of distinctions across the world's languages. The conceptualization of the future is an interesting cognitive phenomenon as we experience the future as a thing of possibility where anything *could* happen. This means it could lend itself to cognitive biases or influences which makes it an interesting topic to study for more abstract operations like linguistic influences. At a glance, it seems reasonable that the experience of time would be a universal phenomenon shared among all human communities. However, this premise was challenged in the work of Benjamin Whorf. Whorf's investigations suggested notable variations in how time is linguistically represented across different languages. His studies were particularly focused on the Hopi language, where he discovered that its speakers described time in a manner distinctly different from Western languages. His observations suggested deep-seated differences in the conceptualization of time, challenging the notion of universal temporal experience. Whorf introduced his hypothesis, the Linguistic Relativity Hypothesis. The hypothesis predicts that language impacts thought in multiple ways, including in the way speakers perceive the world (Whorf, 1956). While there are many criticisms to his work, it points out an interesting potential relationship between language and thought.

Languages can differ vastly in how they grammatically divide and mark time. Some languages have no future tense where they use the same unmarked forms to refer to the present and the future (Dahl, 2000). Others do not distinguish between the present and past, but rather only mark the distinction between experienced and non-experienced (Dahl, 2000). Some languages distinguish precisely between types of futures with one form marking near future events and others marking far further future events (Dahl, 2000). The linguistic relativity

hypothesis holds that differences in how speakers talk about time will cause differences in how they experience it.

The study of linguistic relativity is a large field and there are many ways the linguistic relativity hypothesis has been studied because there are multiple ways that language has been predicted to affect thought (Wolff & Holmes, 2011). Hypothesized influences can be classed into 3 levels: semiotic, structural, and functional (Lucy, 1997). The first level, semiotic relativity, refers to how speaking a natural language can affect thinking in general. This level concerns how language can impact a person's ability to think symbolically. The second level, structural relativity, refers to how speaking one or more different languages can affect thinking. This is more concerned with cross-linguistic differences in morphosyntactic constructions and how those can affect thinking. The last level, functional relativity, refers to how using language in a particular way can affect thinking. This is more about how learning domain specific language aids speakers in understanding and talking about specialized knowledge.

To narrow down the scope of my research, I am specifically studying structural relativity. The main motivation for why I chose to focus on structural relativity is the feasibility of experimentation. A difficulty of studying linguistic relativity is that it can be impractical to study in an experimental setting because it may not be ethical to manipulate the relevant variables. For example, empirically testing semiotic relativity could, in theory, be achieved by restricting a person's access to language either from an early age or by affecting the underlying cortical substrate of language. Operationalizing linguistic relativity with respect to a specific prediction makes testing it more feasible and straightforward. Critically, linguistic relativity holds that the language someone speaks has a potential impact on the way they think. Investigating this

assumption therefore requires examining whether differences in language correspond not only language, but also in non-linguistic cognition.

In the context of the topic of time, a test of the linguistic relativity hypothesis entails investigating whether differences in the way languages grammatically oblige speakers to reference future events impact the way speakers perceive future. With respect to difference in language, I will focus on cross-linguistic differences in the use of obligatory future time reference (FTR) language. Languages differ in the degree to which they require speakers to distinguish the present from the future. A weak FTR language does not obligate speakers to differentiate references to the future from references to the present, whereas a strong FTR language obligates such a distinction. This does not mean that weak FTR languages have no distinctive grammatical expressions for future time, it just means that they are not required to use them.

Work by Chen (2013) suggests how cross-linguistic differences in future reference might impact non-linguistic thinking. Such an impact is potentially revealed in tasks whether people must choose between smaller immediate rewards over larger future rewards. A phenomenon known as psychological discounting occurs when people choose a smaller but more immediate reward over a larger but delayed reward. An example of this is if asked, “Do you want $5 right now or $20 in two weeks?”, that person chooses the $5. Psychological discounting is exhibited in a variety of ways, including savings and health behaviors. Deciding how much money to put into a saving account requires making a decision between spending in the present and not buying something later or spending in the future but being able to buy something of more value. Health behavior requires a similar kind of choice between doing things you want to do in the present that are not health conscious or doing things that are less pleasurable but more health-conscious

for the wake of a healthier in the future. Correlational research done by Chen (2013) has shown a strong relationship between the FTR status of a language someone speaks and their tendency to engage in psychological discounting with respect to savings and health behavior.

The mechanisms by which FTR language status may lead to psychological discounting behaviors have been debated. Chen (2013) has two hypothesized mechanisms: the temporal hypothesis and the precision hypothesis. The temporal hypothesis suggests that FTR language affects future choices by changing how distant future events feel. Strong FTR speakers are expected to have an obligatory distinct future grammar which makes the future events feel more temporally distant. Weak FTR speakers, on the other hand, speak about the future as if it'll happen in the present, so they feel as if the future is temporally closer and imminent. A second hypothesis– the precision hypothesis– suggests that languages with more grammatical time marking would lead speakers to hold more precise beliefs about the timing of events. Because weak FTR speakers do not have to constantly grammatically mark for the future, these speakers do not think precisely about the temporal distinction between the present and future. Future rewards are hypothesized to have more value because of the fuzzy distinction between present and future.

Chen's (2013) temporal distance and prevision hypotheses have recently been challenged by a yet another proposal: The Modal Hypothesis (Robertson, 2022). According to Robertson (2022), the relationship between FTR grammar and psychological discounting is mediated by obligatory certainty FTR which affects speakers' perceptions of certainty. Figure 1 depicts the mechanism underlying the modal hypothesis. The key idea in the modal hypothesis is that speakers of a strong-FTR language are compelled to use future-entailing words such as modal verbs, such as can, *could, might,* and *will*. In a language, such as English, the majority of these

verbs imply a relatively high degree of uncertainty. In the context of a temporal-decision task, relatively high degrees of uncertainty will make future rewards relatively less attractive than present rewards, leading the speakers of a strong-FTR language to engage in more temporal discounting than speakers of a weak-FTR language.

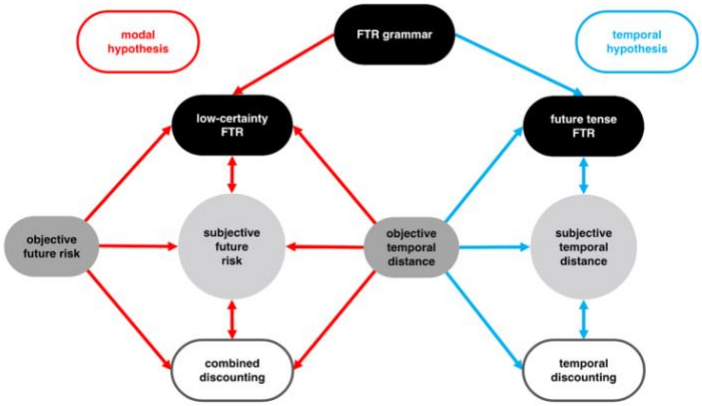


Figure 1: Proposed Mechanisms of Different Hypotheses (Robertson, 2022)

In addition to impacting temporal making, the modal hypothesis suggests that certainty in the occurrence of a future event should be less in the speakers of a strong FTR than in speakers of a weak FTR language. The model hypothesis implies that this tendency to construe future events with different levels of certainty may impact the way people judge the probability of future events, even in situations in which people are asked to assess the probability of a perceptual event. This possibility is examined this thesis, which examines the relationship between FTR grammar, low certainty FTR, and subjective future risk/subjective feelings of certainty in a future event. The non-linguistic cognition I will be examining in this project is event perception and certainty feelings as speakers use their perception of events, they have witnessed to form an opinion about the likelihood of a related future event happening, which they will use to decide how certain they would be that the future event would happen.

I will first provide a literature review to give a summary of important background knowledge on linguistic relativity, future time referencing language, certainty language, and more recent research on structural relativity and psychological discounting (the Linguistics Savings hypothesis and the Modal Hypothesis). I will first explore the work of Benjamin Whorf on Hopi time which inspired the original linguistic relativity hypothesis. Then, I will explore the work of Chen (2013) and Robertson (2022) involving future time referencing (FTR) language and psychological discounting. I will then discuss more technical work on what notions are encoded in future time referencing language and why modal notions are particularly important and prevalent in future time referencing.

Later, I outline my empirical approach and the rationale behind it. The research experiment consists of two parts: a Future-Time Reference Elicitation Task followed immediately by a Subjective Certainty Rating task to establish causal evidence between the FTR language usage and certainty feelings about future events. I then discuss the empirical results and what they mean in relation to my hypotheses. Finally, I will examine potential implications of the study for the fields of linguistic relativity and psycholinguistics and introduce possible speculative topics outside the field of psycholinguistics that could be related to this research.

Literature Review

Linguistic Relativity and Future Time: Benjamin Whorf and the Hopi Time Controversy

The question of whether cross-linguistic differences impact speakers' thoughts was first introduced by Benjamin Whorf. Whorf hypothesized that a speaker's belief about time could be affected by the language they spoke. He studied the Hopi people, an indigenous North American people, and claimed that the Hopi people had no *Western* conceptualization of time (Whorf, 1956). A common misunderstanding of Whorf's claims is that he argued that the Hopi possessed

no concept of time. However, Whorf actually emphasized that the Hopi conceptualized time differently from Western cultures, rather than lacking a temporal framework altogether. In Hopi grammar, the concepts of "past, present, future" are framed in terms of durability or endurance, kinematics (continuous movement through space), or dynamism (the display of force within a process). Unlike in English, where temporal phrases are often lexicalized as nouns, in Hopi, they are expressed as adverbs. This perspective views time not as a discrete entity, but more as an ongoing action or process.

Whorf (1956) posited that the Hopi language does not categorize time into past, present, and future, but rather distinguishes only between non-future and future. This linguistic framework was hypothesized to have an impact on the Hopi's conceptualization of time, leading to a perception of time as a more continuous and forthcoming ("getting later" with the future "approaching"). According to Whorf, this view of time as a continuous flow, where past and present actions accumulate and significantly impact the future, underscores the importance of preparation for what lies ahead. In contrast, the Western conceptualization of time, according to Whorf, was cyclic and segmented, allowing for the division of time into distinct past, present, and future segments, with each day seen as having a clear beginning and end. This perspective makes continuous repetition less meaningful, as it is perceived that "each day is a new day," thereby diminishing the consideration for the long-term consequences of daily actions. Whorf's critique extended to the notion that Westerners may overlook the future implications of their actions, given the compartmentalized view of time where each day stands along.

Despite the intriguing connections Whorf drew between language, thought, and behavior, his work faced criticism for potentially interpreting Hopi culture and language through a Western lens, thus risking an biased characterization of the language. Nonetheless, his exploration into

how linguistic structures might shape perceptions and behaviors remains a compelling aspect of his research.

Linguistic Savings Hypothesis (Keith Chen, 2013)

Chen's (2013) theory on the relationship between cross-linguistic differences and economic outcomes was inspired by literature on future tense typology and observed differences in economic behavior such as savings rate, health behaviors, and retirement assets between countries of different linguistic backgrounds. Inspired by the future tense typology introduced in Dahl (2000), he claimed that FTR grammar impacted those economic behaviors and psychological discounting. His hypothesis was that being required to speak in a *distinct* way about future events leads speakers to take fewer future oriented actions. Grammar that separates the future from the present would make the future feel more distant and saving would be harder. It's important to note that Chen understood the difference between strong and weak FTR language to be the obligation of marking the future using *future tense*.

Another important consideration is that Chen makes these claims in relation to *prediction-based statements*. There are three main types of future statements, and these include: prediction, intention, and scheduling statements. Prediction statements involve referencing things that are out of the control of the speaker and must be guessed from the knowledge about the current state (Dahl, 2000). Intention statements refer to how we plan to act and are restricted to things generally under the control of the speaker (Dahl, 2000). Scheduling statements refer to when we talk about well-known future events for which a specific time is known (Dahl, 2000). Chen specifically defines strong FTR as a language that has obligatory tense marking in prediction clauses. Obligatory FTR is mostly grammaticalized in prediction-based statements and other kinds of tense constructions are allowed in FTR for other kinds of statements (Dahl,

1985, 2000). For example, for a scheduling statement, it is grammatical to say, “I leave at 5 AM tomorrow”. In that case, the present tense is allowed in future referencing because it is a scheduling statement but the same is not true for prediction statements. The motivation to focus on obligatoriness of future tense marking in prediction-based FTR is to focus on a linguistic effect that is grammaticalized and provides something systematic to examine between languages.

Chen used cross-country regressions to show a strong correlation between FTR language status and future-oriented behavior that do not weaken even when other factors are controlled for. He found that speakers of weak FTR languages are more likely to put money into savings, accumulate more wealth by retirement, less likely to smoke, more likely to be physically active, and are less likely to be medically obese (Chen, 2013). These effects were also found when examined within country to suggest that the effect is not due to cultural background. These results were also observed when personal differences like how much people trust others which strengthens the suggested correlation.

While these effects observed from regression analysis were significant and held strong even when controlling for other confounds, a main criticism of his work is that it is only correlational. Because Chen did not experimentally manipulate any variables, the effects observed were never causally proven.

Modal Hypothesis (Cole Robertson, 2022)

As mentioned previously, Chen understood the difference between strong and weak FTR language to be the obligation of marking the future using *future tense*. He defined a strong FTR language as a language that obligates the use *future tense* in referencing the future. His ideas on FTR typology have been criticized for oversimplifying the grammar of prediction-based

statements. Referencing the future involves encoding more than just tense and also encodes notions of certainty because of how the future is a thing of possibility.

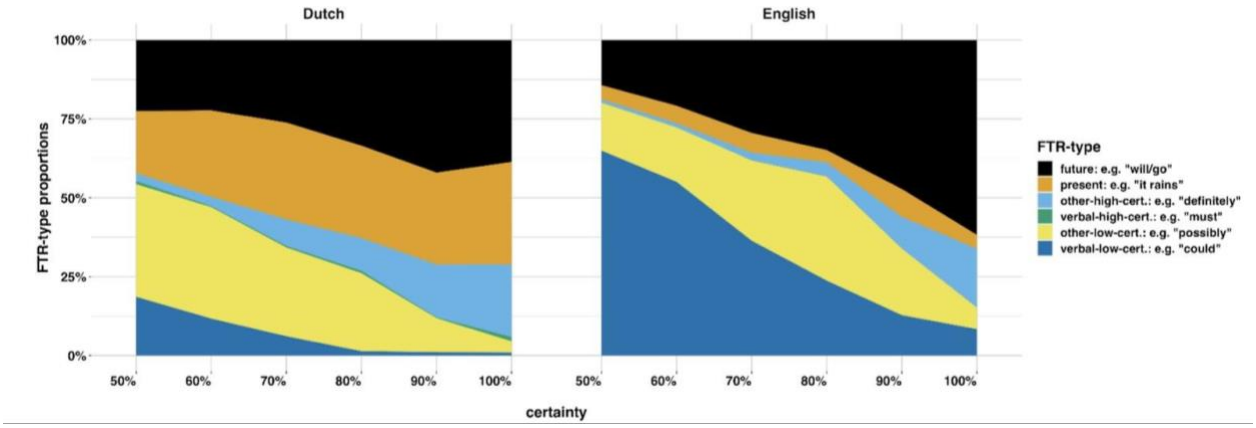


Figure 2: FTR Proportions over Certainty (Robertson, 2022)

Robertson (2022) found that a main difference between English and Dutch speakers was that English speakers tended to use more low-certification modal verbs in FTR and used similar amounts of the future tense compared to Dutch speakers (Figure 1). The obligation to use a low-certification modal verb for strong FTR speakers may cause them to construe future events as more risky and less likely to happen, so they devalue those future events more.

It is a reasonable assumption to make that if speakers are using more low-certainty language, that means they're also more uncertain about what they're talking about. However, even if this is a fair assumption to make, this relationship has not been studied using an experiment, so it has not been causally proven. This lack of research inspired the current research project, and this research project seeks to experimentally establish a potential causal relationship between FTR status, certainty obligatory FTR, and perceptions of the certainty of future events.

The Grammar of Future Time Reference: Tense and Modality

There is an important distinction to be made between the future tense and future time reference. A common misconception is that future time reference is comprised of only the future

tense, but there are other ways to express a future event through certainty-marked language. Future time-referencing language can encode notions of both tense and modality. There is an overlap in the grammatical categories of tense and modality (certainty) but these categories are semantically different. Tense encodes temporal notions and is used to relate the time of the referenced event to the time of the utterance (Lyons, 1968; Mezhevich, 2008). In English and Dutch, future tenses are formed using paraphrastic auxiliary modal verbs that have "go-based" constructions ("will", "going to", "shall").

Modality is used to express degrees of possibility/necessity or obligation. There are four types of modality: deontic, dynamic, bouletic, and epistemic. This schema was taken from *Mood and* Modality (Palmer, 2001), and I will provide a brief description for each of them. Deontic modality refers to expressing what is necessary or expected based on social norms/obligations. Dynamic modality expresses what is possible relative to internal abilities of the entity being referenced. Bouletic modality expresses some state of affair about the desires and hopes of the speaker. Epistemic modality expresses what is probable or certain relative to what the speaker knows. Epistemic modal notions are related the most to notions of risk and prediction examples in psychological discounting so I will mainly focus on epistemic modality for this thesis.

Certainty Language

In English and Dutch, modality is most frequently expressed through a system of modal verbs. For English, these verbs include "can, could, may, might, should, must." In Dutch these verbs include *kunnen 'may'* and *moeten* 'must'. On top of these modal verbs, modality can also be expressed in both English and Dutch through a variety of lexical periphrastic constructions. These constructions can be adverbial or adjectival and examples and in English, include "possibly, possible, definitely, likely".

There is debate about the future tense "will" and if it should be included in the modal category. Because the difference between tense and modality is a semantic one, previous research leans towards considering "will" as a high certainty modal verb (Giannakidou & Mari, 2018; Lyons, 1968) because there are instances where it can be used to express modality without encoding any temporal notion.

Hypotheses

I hypothesize that FTR grammar and perception of certainty are mediated by usage of certainty modal verbs. I propose those variables are related like so:

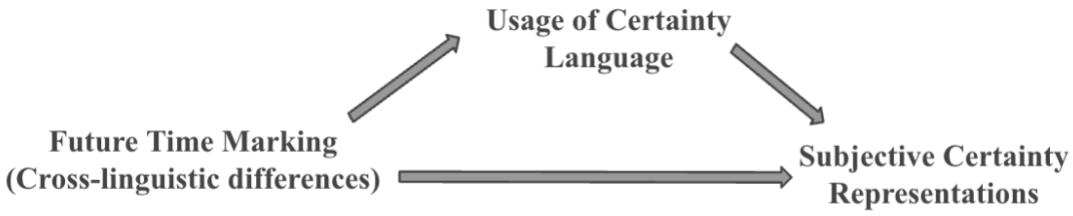


Figure 3: Simple Mediational Model for Variables of Interest

In this experiment, I predict that strong FTR speakers will use more low certainty modal verbs than weak FTR speakers. Because of this, I predict that strong FTR speakers will have be less certain in a future event when asked to indicate their certainty about the event happening. Lastly, I predict the use of the modal verbs will be statistically significant as a mediating variable between FTR grammar and subjective certainty feelings.

Methods

Ethics Approval

An IRB proposal was reviewed and approved by the Emory University IRB. The proposal included consent procedures, consent materials, a GDPR statement, and an approved cultural competence letter. The approval number is 00007221.

Participants

Participants were recruited using an online platform, Prolific. In total, N = 104 participants completed the study.

Dutch and English Participants

Fifty native Dutch speaking participants were recruited. The screening requirements for the Dutch participants were that their first language, earliest language in life, and primary language were Dutch, and that they were 18 years old or older. They were also required to reside in a primarily Dutch-speaking country, either Belgium or the Netherlands (IP addresses that were not from these countries were excluded from participating). Dutch participants balanced for sex. Fifty percent of the participants were male (*n* = 25) and 50% of the participants were female (*n* = 25). Sixty-four percent of the participants were in the 18-29 age group (*n* = 32), 14% of the participants were in the 30-39 age group (*n* = 7), and 22% of the participants were in the 40-49 age group (*n* = 11).

Fifty-four native English-speaking participants were recruited. The screening requirements was that their first language, earliest language in life, and primary language were English, and that they were 18 or over. English participants also matched for sex (50/50), and ages were matched against the Dutch sample. They were also required to reside in the US (by IP filtering). Fifty-six percent of the participants were male (*n* = 30) and 44% of the participants were female (*n* = 24). Sixty-three percent of the participants were in the 18-29 age group (*n* = 34), 11% of the participants were in the 30-39 age group (*n* = 6), and 26% of the participants were in the 40-49 age group (*n* = 14).

Materials and Measures

Qualtrics Survey

The experiment was implemented using the online survey platform Qualtrics. There were two versions of the survey, one completely in English (for English participants) and one completely in Dutch (for Dutch participants). The content, order, length, and formatting of the were otherwise identical.

Part 1: FTR Elicitation Task

To establish FTR usage, an FTR elicitation task was used. This was adapted from Robertson (2022). This was essentially a "fill in the blank" paradigm, where sentences were given with the main verb inflected, and participants were asked to conjugate it "as though speaking to a friend". Some sentences also included some context information, e.g. "A to B", "Doctor to patient", etc. All sentences referred to the future (later on the same day) and all sentences were predictions. Participants were also given some "certainty information" with each question. They were asked to imagine they were "XX% certain" and complete the target sentence as though that were the case. Participants were asked to use all the context information given to them to conjugate the main verb appropriately prompted in the target sentence. The task had one within-subjects factor: certainty information, with five levels (50%, 60%, 70%, 80% or 90%). There were two questions at each level (Appendices H and I), making for a total of 10 trials. Question order was randomized.

Part 2: Subjective Certainty Rating Task

To establish non-linguistic representations of probability, we used a visual paradigm that exposed participants to a known (but withheld) frequency distribution and then asked them to estimate a probability on its basis. For each trial, there was an exposure and test stage. During

the exposure stage, participants viewed 10 images comprising red and blue pixels. Each image auto-advanced after 700ms. At the end of the exposure stage, participants then completed the response phase: They were asked to use a slider to rate the likelihood that a random pixel chosen from an image in the preceding exposure stage would be red. Participants rated the likelihood between "not RED" (0.00) and "RED" (999.99). Numerical values were not shown to participants, and participants were instructed to "try to answer quickly and intuitively, without thinking about it too much". All sliders started out in the middle of the scale (500). Participants could not advance until they had moved the slider (even if just to move it back to the starting position). The task had on within-subjects factor: pixel proportion, the mean proportion of red pixels over the 10 images in each exposure stage. These were matched with the levels of the certainty information in the FTR elicitation task (50%, 60%, 70%, 80%, and 90%). It was not the case that all n=10 images in each trial had the same proportion; rather they were drawn from a normal distribution with a mean of precisely, for example, 50% red pixels(Appendices J-N). The stimuli images were generated using custom python software. Trial order was randomized.

Procedure and Design

Experiments were carried out online and took around 10-12 minutes for the English participants. The Dutch participants took slightly longer and took around 10-20 minutes. Participants had to indicate informed consent and that they were over the age of 18 to participate in the experiment. Before starting the experiment, participants were told to ask for clarification at any point, if necessary, or that they were free to leave the experiment at any point in the study. After, participants filled out demographic questions, entered the FTR elicitation task and then the Subjective Certainty Rating Task. Participants were asked to rate the likelihood that a random dot picked from the previous images would be red. It is important that the subjective certainty

rating task immediately follows the FTR elicitation task to test the presence of a causal effect as the hypothesized cause is immediately before the predicted effect. Throughout the FTR elicitation tasks and the subjective certainty rating tasks, there were attention checks embedded throughout to ensure that participants were answering honestly and paying full attention. There were 2 FTR attention checks where participants were told to type a specific word ("carrot" and "dance") into the response box. There was 1 subjective certainty attention check where participants were told to move the slider all the way to the right (999). If participants failed to answer the attention check correctly, they were rejected from the study and did not receive compensation. Because of the randomization of the block order, some participants saw the subjective certainty rating attention check as their last block after completing all the actual trial blocks. For those participants, even if they failed the attention check they received compensation, and their responses were included in the data analysis. Participants were compensated at $12/hour for completing the full experiment or as in the case mentioned above.

The design of this study was a cross-sectional quasi-experimental design. The experiment did aim to establish a causal effect between variables, but participants were not randomly selected and were recruited based on nonrandom criteria. The independent variable in this experiment was FTR status, the mediating variable was the usage of low-certainty modal verbs, and the dependent variable was certainty ratings of future events. The experiment also used a between-subjects design as two different participant populations were tested simultaneously.

Results

Data Preparation

Any incomplete data was removed from the data set. Trials where the participants finished the entire survey or only missed the attention check when it was the last trial block, they would see were all included in the data set.

Qualitative data was recoded into numeric values for data analysis. Participant's native language was recoded using the values 1 and 0. English was assigned a value of 1 while Dutch was assigned a value of 0.

Data from the Dutch survey responses and English survey responses were merged into one data set. In this master data set, responses were separated by certainty level. The mean usage of low certainty modal verbs was calculated for both Dutch and English speakers. The mean certainty response at each certainty level for Dutch and English speakers was also calculated. To control for potential recency effects, the average certainty response level across all certainty levels was also calculated for Dutch and English speakers.

FTR Elicitation Data

Responses from the FTR elicitation task were recoded using values 1 and 0. In particular, we used the FTR-classifier (Robertson and Roberts, 2023) to code (1) responses that used low-certainty modal verbs ("could, may, might, should, can sometimes"). We refer to this as "verbal low-certainty".

Each certainty level in the FTR task had 2 trials associated with it so we calculated the mean verbal low-certainty over for each participant and certainty level.

Subjective Certainty Rating Task

Data downloaded directly from Qualtrics had a numeric value for the position that participants put the slider at. This number ranged from 0 to 999.99. Data was examined to see if participants were all answering in good faith. Participants who had the same number for all trials, meaning that they had been putting the slider in the same position, were considered to be responses that were answered not earnestly and excluded from the data set. Some responses alternated between 0 and 999 throughout all trials and those responses were also excluded as it suggested that those participants had been putting the slider in the easiest positions without the participants genuinely considering the response. In total, 3 responses were excluded from data analysis. Two responses from the English data set were excluded and 1 response from the Dutch data set was excluded.

Overall Descriptives of Measures

The means for the average use of verbal low certainty verbs for both English and Dutch speakers are presented in Table 1. The mean usage of verbal low-certainty verbs for Dutch speakers was 0.09 while the mean usage of verbal low-certainty verbs for English speakers was 0.31.

Table 1: *Overall descriptive measures for low certainty modal verb usage*

	Dutch	**English**
Low Certainty Modal Verb Usage	0.09	0.31
Mean Difference	0.22	

The mean difference in use of verbal low certainty verbs between Dutch and English speakers was 0.22 (Table 1) and this was a significant mean difference (Figure 4). On average, English speakers used higher proportions of verbal low-certainty verbs than Dutch speakers (Figure 4).

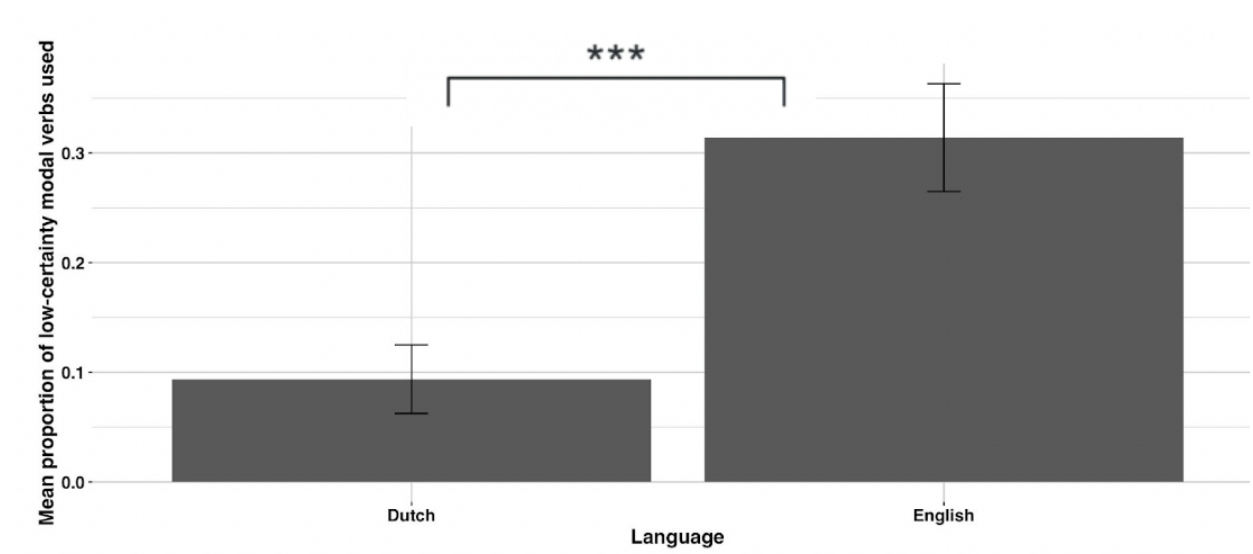


Figure 4: Mean Proportions of Low Certainty Modal Verbs across Dutch and English speakers

The means and standard deviations for the average certainty response at each certainty levels for both English and Dutch are presented in Table 2.

Table 2: *Overall descriptive measures for certainty responses at each certainty level*

	Dutch		**English**	
	M	*SD*	*M*	*SD*
50%	389.29	222.47	398.84	252.27
60%	563.18	200.36	528.68	239.11
70%	664.44	197.29	677.94	213.71
80%	853.23	123.26	824.36	192.40
90%	919.32	81.51	882.96	194.95

Test of Hypotheses

As the frequency of the red dots increased, so did the average certainty responses (Figure 5). A Repeated Measures ANOVA determined that average certainty responses varied significantly across certainty levels for English speakers, $F(4, 220) = 45.50$, $p < .001$. A post hoc analysis showed that while there was not a significant difference between average certainty responses at the 80% level ($M = 824.36$, $SD = 192.40$) and 90% ($M = 882.96$, $SD = 194.95$) but

the average certainty responses at the 50% level (*M* = 398.84, *SD* = 252.27), at the 60% level (*M* = 528.68, *SD* = 239.11), and at the 70% level (*M* = 677.94, *SD* = 213.71) all differed significantly. The Repeated Measures ANOVA determined that average certainty responses varied significantly across certainty levels for Dutch speakers also, $F(4, 200) = 112.42$, $p < .001$. A post hoc analysis showed that the average certainty responses at the 50% level (*M* = 398.84, *SD* = 252.27), at the 60% level (*M* = 528.68, *SD* = 239.11), at the 70% level (*M* = 677.94, *SD* = 213.71), at 80% level (*M* = 677.94, *SD* = 213.71), and at 90% level (*M* = 677.94, *SD* = 213.71) all differed significantly. As the frequency of the red dots increased, speakers felt that the likelihood of a random dot being picked would be red. There was an observed relationship between average certainty responses and dot frequency, but there was no observed cross-linguistic difference as Dutch and English speakers did not vary significantly at each dot frequency (Figure 5).

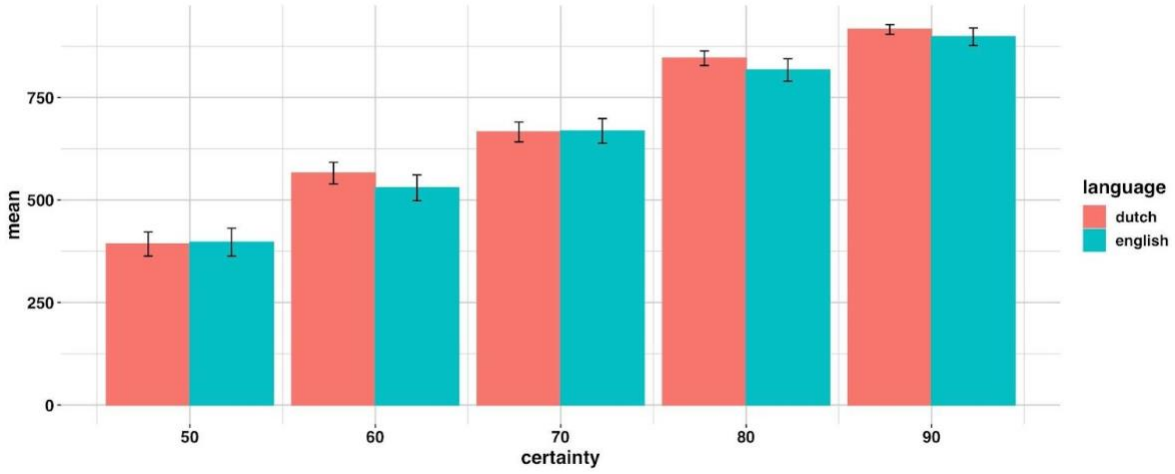


Figure 5: Mean Certainty Ratings across Certainty Levels for Dutch and English speakers

To examine the use of verbal low-certainty verbs as a mediating variable between FTR grammar and certainty ratings, a simple mediation analysis was run with a *p-value* set at .05 with a confidence interval of 90%. A confidence level of 90% was used because the hypothesis a

directional hypothesis that is one-tailed. The hypothesis predicts that English speakers will use more verbal low-certainty verbs which will cause lower ratings of certainty of a future event. Bootstrapping was used in the mediation analysis. In mediation analysis, bootstrapping involves resampling the data with replacement to generate multiple bootstrap samples. This process allows for the estimation of indirect effects and their corresponding confidence intervals, even in cases where the sample size is small, or the distribution of the data is non-normal. By iteratively resampling the data, bootstrapping provides robust estimates of indirect effects, enhancing the reliability and validity of mediation analyses. It is particularly useful when testing complex mediation models or when the assumption of normality is not met, ensuring more accurate inferences about the relationships among variables in a model.

Important values from the mediation analysis are reported in Table 3. As a reminder, an indirect effect suggests a causal hypothesis where an independent variable causes a mediating variable which then causes a dependent variable.

Table 3: *Mediation Analysis for Outcome Variables*

Path	**Outcome Variable**				**90% Confidence Interval**		
		Coefficient	*SE*	*t*	Lower	Upper	p
FTR grammar -> Usage of Low Certainty Modals	Verbal Low Certainty Verbs	0.214	0.042	5.16	0.146	0.2846	.0000
Usage of Low Certainty Modals-> Subjective Certainty Ratings	Subjective Certainty Responses	-44.20	54.6	-.80	-134.84	46.43	.4201

Table 4: *Mediation Analysis for Direct and Indirect Effects*

Effect	**90% Confidence Interval**				
	Coefficient	*SE*	*t*	Lower	Upper
Direct	1.718	25.8	0.07	-41.16	44.59
	Coefficient	*BootSE*		BootLower	BootUpper
Indirect	-9.52	0.042	5.16	0.146	0.2846

There was a significant positive effect of FTR grammar on the usage of verbal low-certainty verbs, a = 0.2154 (90% *CI:* .1462, .2846), p = .0000. The closer the FTR grammar was closer to English FTR grammar, there more usage of verbal low certainty verbs was observed. There was an insignificant negative effect of the usage of verbal low-certainty verbs on certainty ratings, b = -44.2013 (90% *CI:* -134.83, 46.44), p = .4201. While this result is statistically insignificant, the correlation coefficient goes in the predicted direction according to the hypothesis. The hypothesis was that the usage of verbal low certainty verbs would decrease the subjective certainty ratings. Although insignificant, the results suggest there is a negative correlation and as the usage of verbal low certainty verbs increase, the subjective certainty ratings decrease (Figure 6).

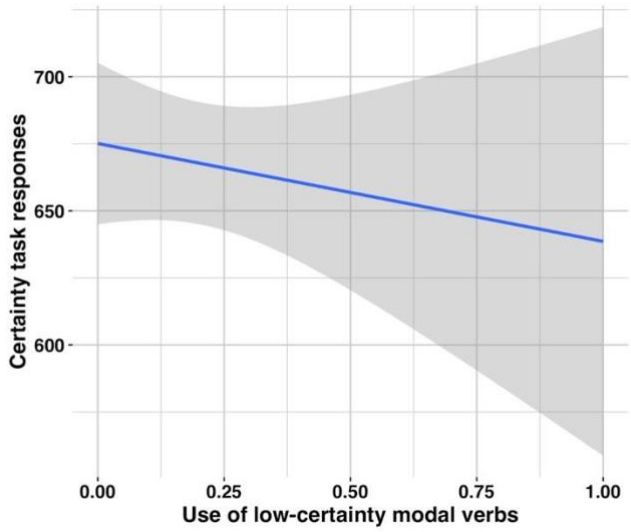


Figure 6: Correlation between Certainty Responses and Usage of Low Certainty Modal Verbs

There was an insignificant negative indirect effect of FTR grammar on certainty ratings, *c'* = -9.5212 (90% *CI:* -27.92, 6.41). While this result is statistically insignificant, the correlation coefficient goes in the predicted direction according to the hypothesis. The hypothesis was that strong FTR language speakers would have decreased subjective certainty ratings. Although insignificant, the results suggest there is a negative correlation and stronger the FTR status, the less subjective certainty ratings of a future event.

There was an insignificant positive direct effect of FTR grammar on the certainty ratings, *c* = 1.718 (90% *CI*: -41.16, 44.59), *p* = 0.94. The correlation coefficient goes in a direction which is unpredicted from the hypothesis. The hypothesis was that weaker FTR language speakers would have decreased subjective certainty ratings. Although insignificant, the results suggest that as the FTR status becomes stronger, the certainty ratings increase.

The total effect refers to the overall association between an independent variable (IV) and a dependent variable (DV) without considering any intervening variables (mediators). The total effect is calculated by summing the indirect and the direct together. The total effect from the

meditation analysis was -7.8032. This means that without considering the intervening variables, overall the effect has a negative correlation between FTR status and certainty responses. This is negative direction is expected from the Modal hypothesis as it suggests that the stronger the FTR language status, the less certain speakers feel about a future event.

Discussion

The current study examined the relationship between FTR grammar and certainty in future events with usage of low certainty modal verbs as a potential mediating variable. This experiment sought to establish a causal effect between all three variables. There were three hypotheses that were examined in this experiment: 1) strong FTR speakers would use more low certainty modal verbs in comparison to weak FTR speakers, 2) strong FTR speakers would have lower certainty ratings in comparison to strong FTR speakers, and 3) the usage of low certainty modal verbs would be a mediating variable between FTR grammar and certainty ratings.

This study is important because it adds to the gap in literature about certainty and the modal hypothesis and would provide more insight on a possible mechanism for the observed relationship between FTR language status and psychological discounting behavior. Previous hypotheses disregard modality as an important aspect of future time referencing and choose to focus mainly on tense as a main notion of future time referencing language. Previous studies have suggested that FTR grammar affects psychological discounting through encoding notions of certainty and possibility (Robertson, 2022) and this study seeks to provide support for this hypothesis. Considerations from this hypothesis would add to the field of FTR grammar and future oriented behavior which have practical implications as it could provide insight on decision-making in important situations like savings and retirement habits or general health

behavior. Considerations from this hypothesis would also add to the field of FTR typology and provide insight on if modality is an important encoding of FTR language.

Interpretation of Mean Usage of Low Certainty Modal Verbs

English speakers had a higher average usage of low certainty modal verbs in comparison to Dutch speakers (Figure 4) and the means were significantly different as previously mentioned. This provides support for the first prediction made in this thesis. This finding provides support for the claim that strong FTR language speakers use more low certainty modal verbs and validates previous studies which have found the same results. Because the future is something of possibility and it technically never is experienced, obligatory future time marking usually involves obligatory encoded notions of possibility, intention, certainty, obligation, desire, necessity, conditionality, chance, and probability (Palmer, 2001).

Interpretations of Average Certainty Response by Certainty Level

For both the English and Dutch participants, as the certainty level changed the average rating in certainty also changed (Figure 5). As the associated certainty level increased, the average certainty response also increased. As confirmed by the Repeated Measures ANOVA analysis, the differences of the average certainty responses between certainty levels were significantly different between all levels for both the Dutch and English speakers except between levels 80% and 90% for English speakers. These results are important because they validate the paradigm of the subjective certainty rating task. As the mean proportion of red dots changed, so did the participants' certainty rating of a random dot that is picked being red. This means it was correct to assuming that participants would use their perception of previous events to estimate a likelihood for a future event. This means that the experiment was successful in measuring certainty in relation to event perception, a non-linguistic cognition, and that our manipulation to

measure uncertainty was appropriate. Because the hypothesized pathway is that FTR grammar affects the non-linguistic perception of events and this affects the feelings of certainty of the speakers, these results validate the latter part of the hypothesized pathway.

Interpretations of the Relationship between Usage of Low Certainty Modal Verbs and Certainty Responses

Looking solely at the usage of low certainty modal verbs and certainty responses shows that there is a negative correlation between the two variables (Figure 6). This means that the more low certainty modal verbs that are used, the lower the certainty responses become. Although this correlation is insignificant, the direction of the correlation is as expected. The negative correlation between certainty responses and the usage of low certainty modal verbs is a promising result as there is some slight correlation. The insignificance of this result does not necessarily mean that there is no effect observed, it could possibly speak more to the power of the study. The study was not conducted with a substantially large sample size which could suggest that the study is underpowered and that is why an insignificant result has been observed.

Interpretations of the Relationship between Usage of Low Certainty Modal Verbs and Certainty Responses

The results from the simple mediation analysis show that there is a significant indirect effect in the positive direction of FTR status on the usage of low certainty modal verbs, but the indirect effect of the usage of low certainty modal verbs on subjective certainty ratings, the direct effect of FTR status on subjective certainty ratings, and the total effect are all insignificant. The indirect effect of the usage of low certainty modal verbs on subjective certainty ratings was in the negative direction and this provides a promising result as that is what is expected. It would be expected that the more low certainty modals verbs that are used, the lower the subjective

certainty responses become. The insignificant result may be due to the study being underpowered because of the smaller sample size and not necessarily because there is no effect. The observation that the indirect also goes in the expected direction also supports this possibility. The direct effect goes in the negative direction. The lower the FTR status value, the higher the subjective ratings of certainty. In the data recoding, English was recoded to a value of 1 while Dutch was recoded to a value of 0. A lower value in the FTR status variable means that it is closer to Dutch FTR status while a higher value in the FTR status variable means that it is closer to English FTR status. This means the closer to Dutch FTR status, the higher the subjective ratings of certainty. This aligns with the original prediction that Dutch speakers, who have been recoded to a smaller numeric value than the English speakers, would be more certain in their ratings of future events and this causes them to not devalue those future events as much (Robertson, 2022). All of these effects being aligned the correct direction provides more support for the possible presence of an effect but just an underpowered study sample.

However, considering the possibility that the lack of significant effects is not due to an underpowered sample but rather there is no effect, this could possibly be due to multiple reasons. Firstly, it provides support against the mechanism of the modal hypothesis and suggests that certainty may not be a mediating variable. This does not completely discount the modal hypothesis as modality encodes many notions other than certainty (Palmer, 2001). This could possibly suggest that another notion in modality that is closely related to certainty may be a mediating variable. Another consideration is that the expected effect did penetrate at the perception level but was overridden by other cognitive effects at higher levels of cognition when the participants are producing the response. Some suggest that linguistic representations are recruited in the production of externalized responses (Kompa and Mueller, 2020), and this

could've interfered with the expected effect of FTR language and low certainty modal verbs in the perception level. An important consideration is that a lot of the Dutch speakers were bilingual in Dutch and English because it is hard to find a lot of Dutch speakers who are not L2 speakers of English. The Dutch speakers' proficiency in English could've had an offline effect in the higher levels of cognition when linguistic representations are recruited in producing a response and interfered with the online linguistic effect that was expected to be seen in non-linguistic cognition.

Limitations

There are several limitations associated with the design of the study. Firstly, this study was carried out online which left for a lot of external influences that could not be controlled for. Participants took the study in varying environment, and this could've influenced their performance. Although attention checks were embedded in the study to ensure participants were paying attention throughout the study, this did not guarantee their full and earnest effort in completing the study.

The sample size is also a limitation of the study. Not only is it on the smaller size that it possibly underpowered the study, but the sample size is also too small to generalize the results to the larger population of the participants that were being observed. Measures were taken for the sample to be representative in sex ratios, but other important demographics were not matched to be representative of the bigger participant populations, so generalizability is also a limitation of this study.

Because of the experimental design of this study, this study has high internal validity but low external validity because of the structuredness and artificiality of the study. Because of this, the observed results have low generalizability to other natural settings and situations.

A general limitation of studying structural relativity in an experimental setting, as mentioned before, is that there is no way to measure completely non-linguistic cognition because the responses must be communicated from participant to research or the prompts have to be communicated to from researcher to participant in some way. This research may not be able truly test structural relativity or observe true structural relativity effects because of this practical constraint.

Implications

These study results have real world contributions as they can be applied in the larger scope of cross-linguistic research and future oriented behaviors. Not only are they important for the discussion of how economic behaviors can vary between countries and how language play a part in it, but it also has important contributions to global health behaviors. Because health behaviors also involve psychological discounting, the study of how language affects or doesn't affect it is an important consideration because it could suggest potential linguistically mediated health interventions. This research also has implications for the larger field of linguistic relativity and provides insight for the debate between linguistic relativity and linguistic universalism. If there are actually no observed effects from this study, then it provides support for a linguistic universalism perspective. However, if further analysis shows support for the effects that were predicted to be observed, then it can add more nuanced considerations to the study of linguistic relativity and provide for support for the linguistic relativity hypothesis.

Future Directions

Firstly, because the lack of significant results in this study could potentially be attributed to the small sample size of the participants, this study should be replicated with a larger sample size to increase statistical power. This study should also be replicated with a more representative

participant population of Dutch and English speakers. If possible, this study should be replicated so that the only Dutch speakers who participate are monolingual in Dutch.

Future research should expand this project to include a later step of psychological discounting as that is the next step in the proposed mechanism. To relate the research done in this project in a causal way to psychological discounting effects, an additional step should be added to test if the same previously observed psychological discounting behaviors are also observed.

References

Chen, K. (2013). The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. American Economic Review, 103(2), 690–731. https://doi.org/10.1257/aer.103.2.690

Dahl, Östen. "The grammar of future time reference in European languages." Tense and Aspect in the Languages of Europe, 6 June 2000, pp. 309–328, https://doi.org/10.1515/9783110197099.2.309.

Giannakidou, A., & Mari, A. (2018). A unified analysis of the future as epistemic modality: The view from Greek and Italian. Natural Language and Linguistic Theory, 36(1), 85–129. https://doi.org/10.1007/s11049-017-9366-z

Kompa, N. A., & Mueller, J. L. (2020). How Abstract (Non-embodied) Linguistic Representations Augment Cognitive Control. Frontiers in psychology, 11, 1597. https://doi.org/10.3389/fpsyg.2020.01597

Lucy, J. A. (1997). Linguistic relativity. Annual Review of Anthropology, 26, 291–312.

Lyons, J. (1968). Grammatical categories. In Introduction to theoretical linguistics (pp. 270–333). Cambridge: Cambridge University Press

Mezhevich, I. (2008). A Time-relational approach to tense and mood. In Proceedings of the West Coast conference on formal linguistics (Vol. 27, pp. 326–334). Somerville, MA: Cascadilla Proceedings Project.

Palmer, R. F. (2001). Mood and modality (2nd ed.). Cambridge: Cambridge University Press

Robertson, C. (2022). *The future is a thing of possibility: Cross linguistic differences in modal time reference affect (risky) intertemporal decisions* [Unpublished Doctoral Dissertation]. Oxford University, Radboud University.

Robertson, C., & Roberts, S. G. (2023). Not when but whether: Modality and future time reference in English and Dutch. Cognitive Science, 47(1). https://doi.org/10.1111/cogs.13224

Whorf, B. (1956). Language, thought, and reality: Selected writings of Benjamin Lee Whorf (J. B. Carroll, Ed.). Cambridge, MA: MIT Press.

Additional Materials

Appendices

Appendix A

Prolific Study Notice (English)

This study is being done to further investigate previous theories on linguistic relativity and decision making. You are being asked to be in this research study because we are seeking native English speakers who reside in the US. This study will take 10-12 minutes and you will be compensated at a rate of $12/hour.

Appendix B

Prolific Study Notice (Dutch)

Dit onderzoek wordt gedaan om eerdere theorieën over linguïstische relativiteit en besluitvorming verder te onderzoeken. Je wordt gevraagd deel te nemen aan dit onderzoek omdat we op zoek zijn naar moedertaalsprekers van het Nederlands die in landen wonen waar de meerderheid van de bevolking Nederlands spreekt (België of Nederland). Dit onderzoek duurt 10-12 minuten en je krijgt $12 per uur.

Appendix C

Informed Consent Question (English)

Please tick the boxes below to indicate you consent to take part in the study.

The purpose of this study is to look at how language can affect non-linguistic behavior. The experiment is being carried out by Tammy Lu, undergraduate researcher in the Phillip Wolff Language Biomarker Group at Emory University.

- ☐ I confirm that I have read and understand the information sheet. I have had the opportunity to consider the information, ask questions and have these answered satisfactorily.
- ☐ I understand that my participation is voluntary and that I am free to withdraw at any time during data collection, without giving any reason and without affecting my legal rights, but that because the data are anonymous I will not be able to withdraw after the collection of my data is complete.
- ☐ I understand that this project has been reviewed by the Emory University Research Ethics committee and has received ethics clearance.
- ☐ I understand that all data collected will remain completely confidential and that the maintenance of confidentiality of this information is subject to normal legal requirements and that all data may be kept past the term of this project in fully anonymised form, on password protected computers and, when not digital, under lock and key.
- ☐ I understand that anonymised data collected during the study may be looked at by responsible individuals where it is relevant to my taking part in this research. The anonymous data may be used in research publications, and a student honors defense.
- ☐ I understand how to raise a concern or make a complaint.
- ☐ I am 18 years of age or over.
- ☐ Given the above, I agree to take part in this study.

Appendix D

Informed Consent Question (Dutch)

Het doel van deze studie is om te kijken hoe mensen de interne structuur van gebeurtenissen waarnemen. Het experiment wordt uitgevoerd door Cole Robertson, afgestudeerd onderzoeksassistent bij de Social and Evolutionary Neuroscience Research Group aan de Universiteit van Oxford, en promovendus aan de Max Plank International School for Language Sciences, Radboud University. Vink de onderstaande vakjes aan om uw toestemming aan te geven.

- ☐ Ik bevestig dat ik de informatie gelezen en begrepen heb. Ik heb de gelegenheid gehad om de informatie te overwegen en vragen te stellen, welke naar tevredenheid zijn beantwoord.
- ☐ Ik begrijp dat mijn deelname vrijwillig is en dat ik vrij ben om me op elk moment tijdens het experiment terug te trekken , zonder enige reden te hoeven geven en zonder dat het gevolgen heeft voor mijn wettelijke rechten, omdat de data anoniem is. Ik kan me niet meer terugtrekken nadat het verzamelen van data compleet is.
- ☐ Ik begrijp dat dit project beoordeeld is door de Universiteit van Oxford Central University Research Ethics Committee, en ethische goedkeuring heeft ontvangen.
- ☐ Ik begrijp dat alle verzamelde data volledig vertrouwelijk blijven en dat het behouden van de vertrouwelijkheid van deze informatie onderhevig is aan de gebruikelijke wettelijke eisen. Ik begrijp dat alle data ook na dit experiment volledig geanonimiseerd bewaard zullen blijven, op wachtwoord beschermde computers of, wanneer deze niet digitaal zijn, achter slot en grendel.
- ☐ Ik begrijp dat de geanonimiseerde data die tijdens de studie verzameld is, bekeken wordt door verantwoordelijke individuen van de Universiteit van Oxford en de Radboud Universiteit, waar het relevant is voor mijn deelname aan deze studie. De geanonimiseerde data mag gebruikt worden in onderzoekspublicatiesen in een proefschrift.
- ☐ Ik begrijp hoe ik mijn bezorgdheid kan aankaarten en een klacht kan indienen.
- ☐ Ik ben 18 of ouder.
- ☐ Gezien het bovenstaande, geef ik mijn toestemming om deel te nemen in deze studie.

Appendix E

Demographics Questions (English)

1. What is your prolific ID?
2. Is English your first language?
3. Do you speak any other languages apart from English?
4. **Select (up to) your top three second languages excluding English:**

   a. First most fluent:
   b. Second most fluent:
   c. Third most fluent:
5. **Rate your fluency between 1 and 5 for your selected second languages using the following scale:**
6. Please enter your age in numbers (e.g. 57)
7. Please enter your sex.

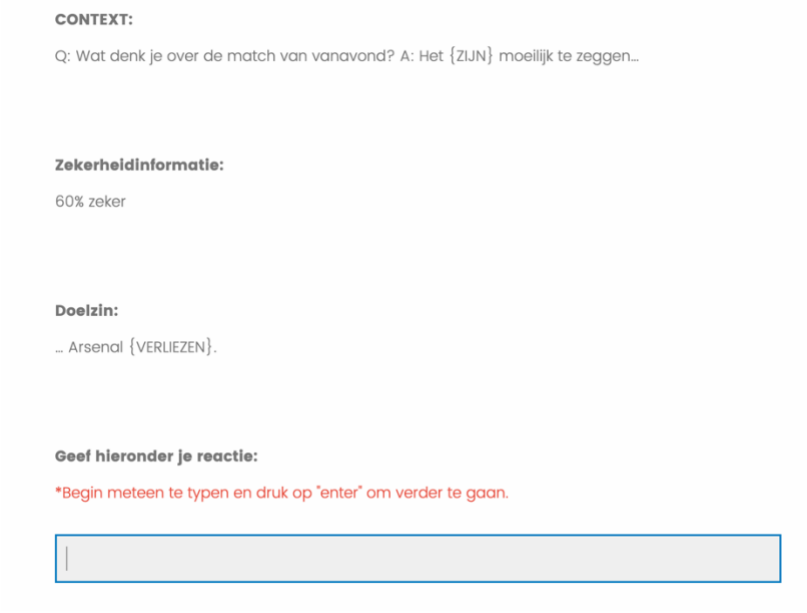
Appendix F

Demographics Questions (Dutch)

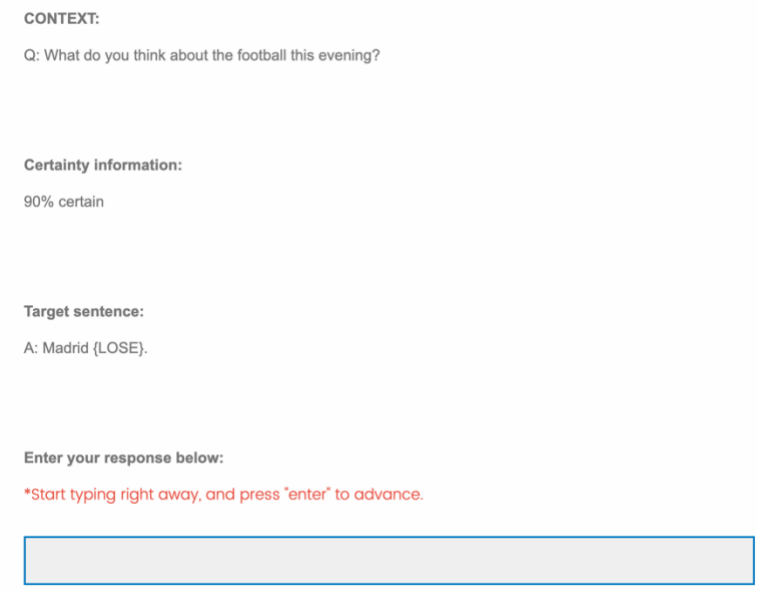
1. Wat is uw prolific ID?
2. Is Nederlands je eerste taal?
3. Spreek je naast Nederlands nog andere talen?
4. **Maak een lijst van uw top drie tweede talen (behalve het Nederlands):**
   a. Meest vloeiend:
   b. Tweede meest vloeiend:
   c. Derde meest vloeiend:
5. **Beoordeel uw vloeiendheid tussen 1 en 5 voor uw geselecteerde tweede talen met behulp van de volgende schaal:**
6. Vul je leeftijd in in cijfers (bijv 57)
7. Wat is je geslacht?

Appendix G

FTR Elicitation Task



Dutch



English

Appendix H

FTR Questions (English)

50% certainty:

1. [A to B: Don't you think we should study the extra material for this evening's test...] …there {BE} a question on it.
2. [Q: Should I bring and extra warm jacket? A: It be your decision...] ...it {SNOW} this afternoon.

60% certainty:

1. [Q: Should I bring my umbrella? A: Yes…] …a storm {BLOW IN} this afternoon.
2. [A to B: We should study chapter three tonight...] …there {BE} a question on it.

70% certainty:

1. [A to B: We should study chapter three tonight...] …there {BE} a question on it.
2. [Speaking to a friend who fancies Claire: You should come to the party tonight...] …Claire {BE} there.

80% certainty:

1. [A to B: Bring an umbrella to Germany later today...] …it {RAIN} in Berlin.
2. [Q: What do you think about the football this evening?] A: Madrid {LOSE}.

90% certainty:

1. [A to B: We should definitely study the digestive system for the exam this afternoon...] …there {BE} a question on it.
2. [Q: Should I being my sun hat to the beach? A: Yes...] …it {BE} very hot this afternoon.

Appendix I

FTR Questions (Dutch)

50% certainty:

1. [A to B: Don't you think we should study the extra material for this evening's test...] …there {BE} a question on it.
2. [Q: Should I bring and extra warm jacket? A: It be your decision...] ...it {SNOW} this afternoon.

60% certainty:

1. [A to B: I would bring your bathing suit this afternoon...] …the tempurature {HIT} 45 degrees.
2. [Q: What do you think about the match tonight? A: It {BE} hard to say…] …Arsenal {LOSE}.

70% certainty:

1. [A to B: Bring some sun cream this afternoon...] …it {BE} sunny!
2. [Q: Should I bring my umbrella? A: Yes…] …a storm {BLOW IN} this afternoon.

80% certainty:

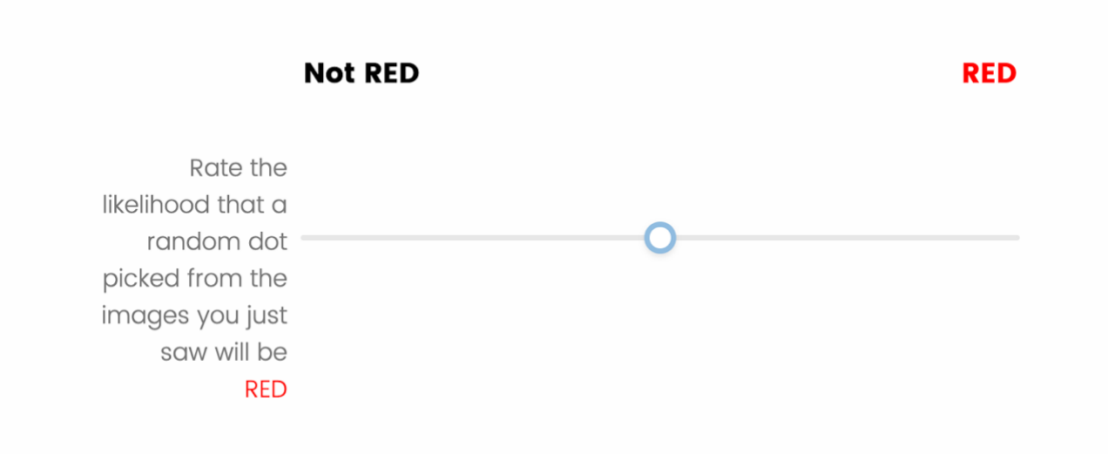
1. [A to B: We should study chapter three tonight...] …there {BE} a question on it.
2. [Speaking to a friend who fancies Claire: You should come to the party tonight...] …Claire {BE} there.

90% certainty:

1. [A to B: Bring an umbrella to Germany later today...] …it {RAIN} in Berlin.
2. [Q: What do you think about the football this evening?] A: Madrid {LOSE}.

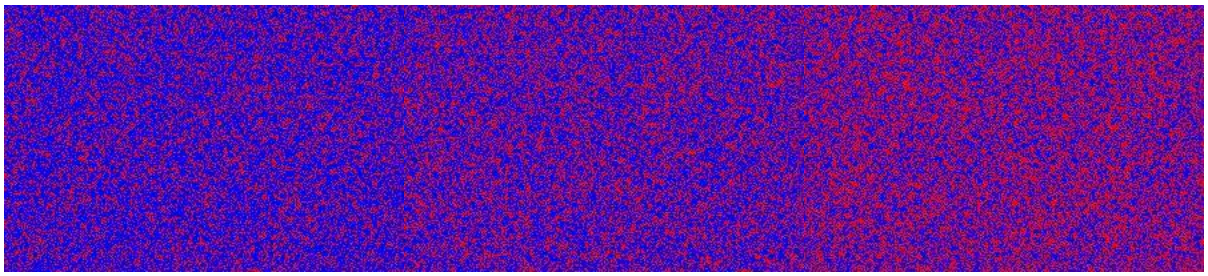
Appendix J

Subjective Certainty Rating Task

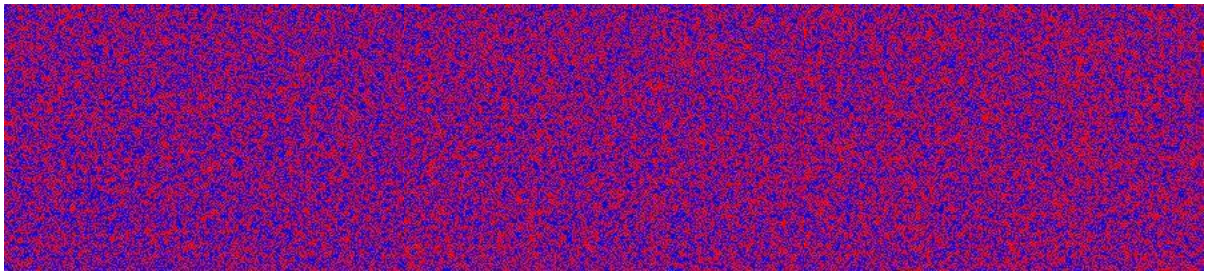


Appendix K

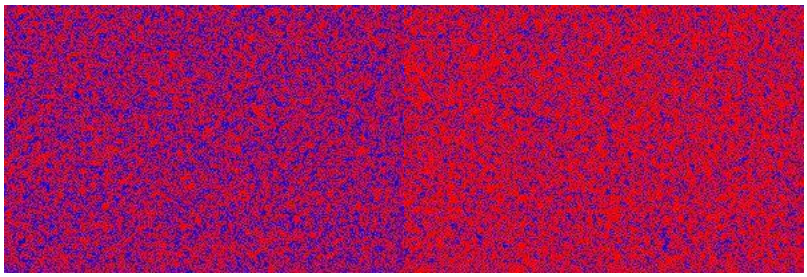
Red/Blue Dot Stimuli (50% Red)



31% 37% 46%



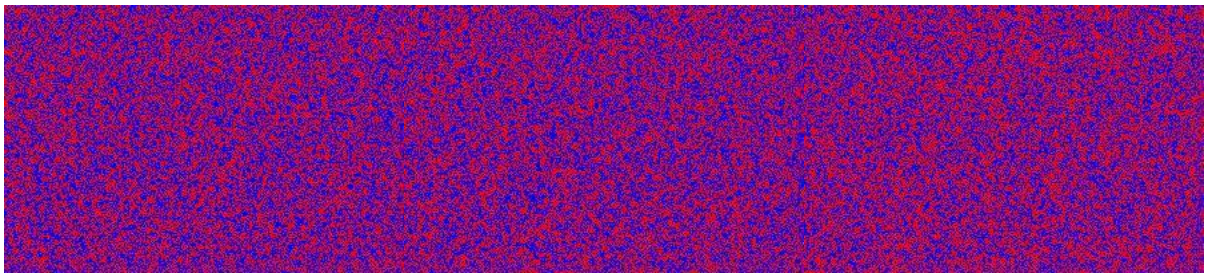
49% 51% 52%



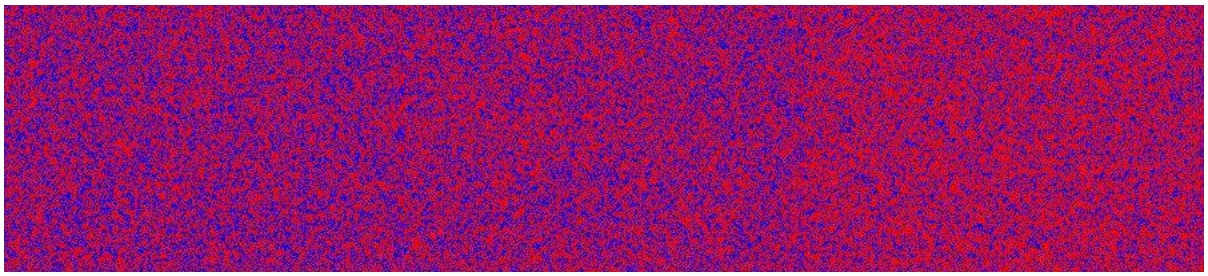
60% 71%

Appendix L

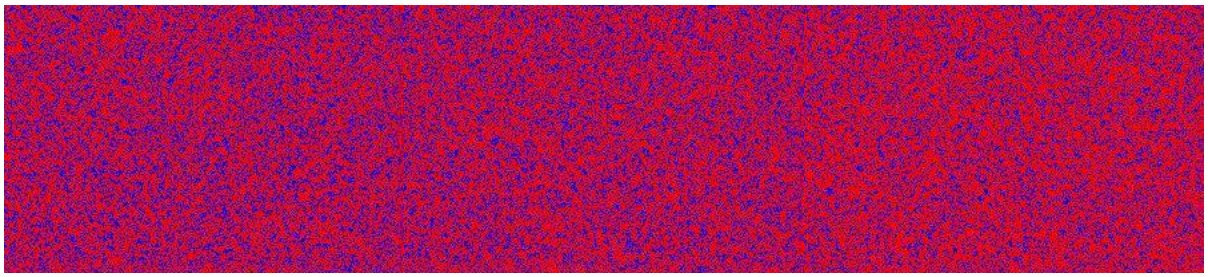
Red/Blue Dot Stimuli (60% Red)



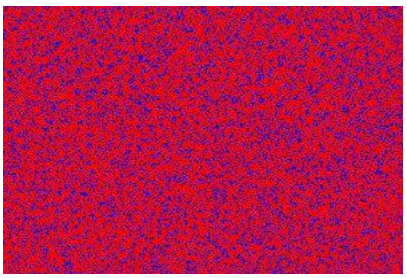
49% 50% 52%



56% 57% 64%



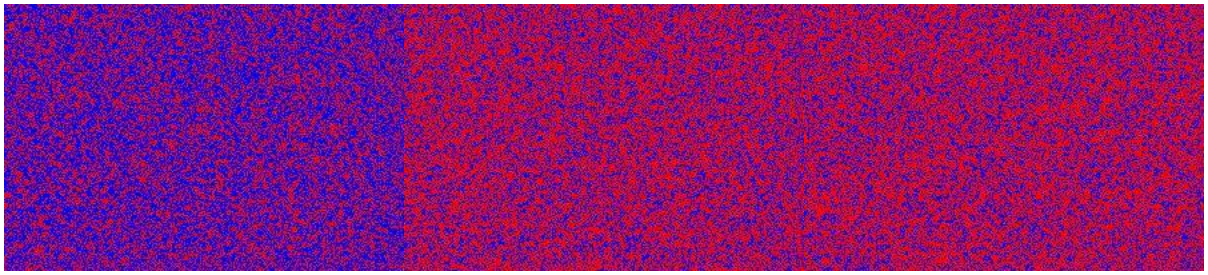
65% 66% 69%



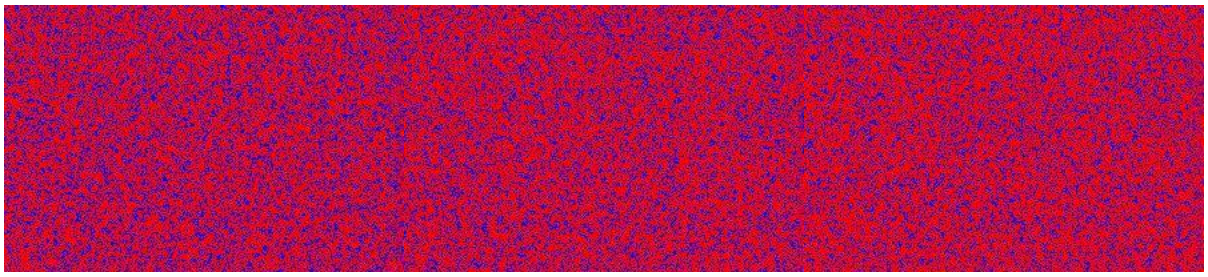
72%

Appendix M

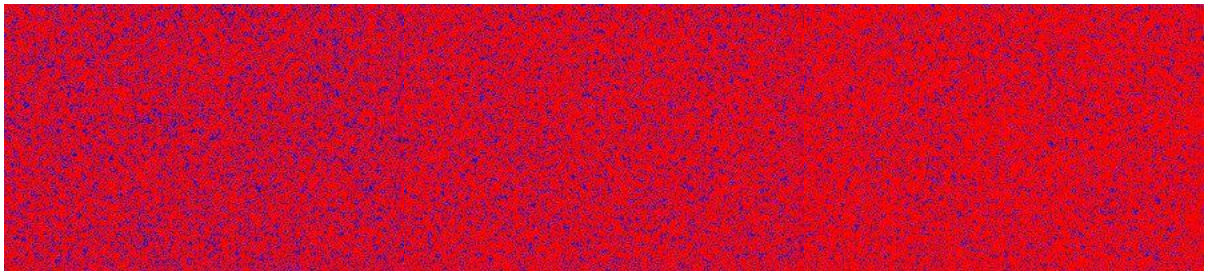
Red/Blue Dot Stimuli (70% Red)



42% 60% 62%



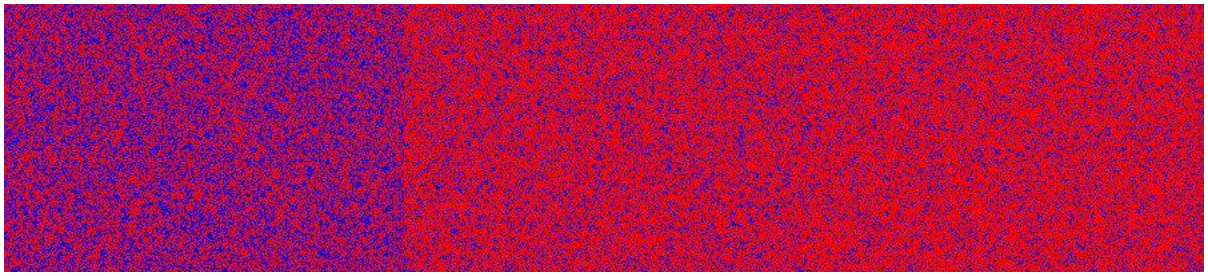
69% 71% 72%



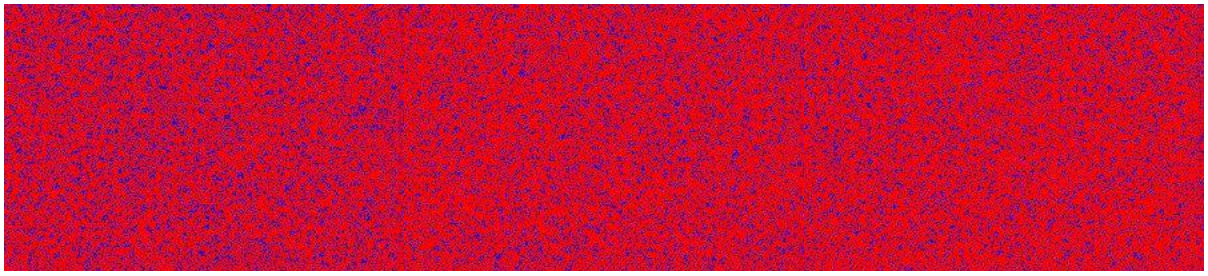
78% 81% 85%

Appendix N

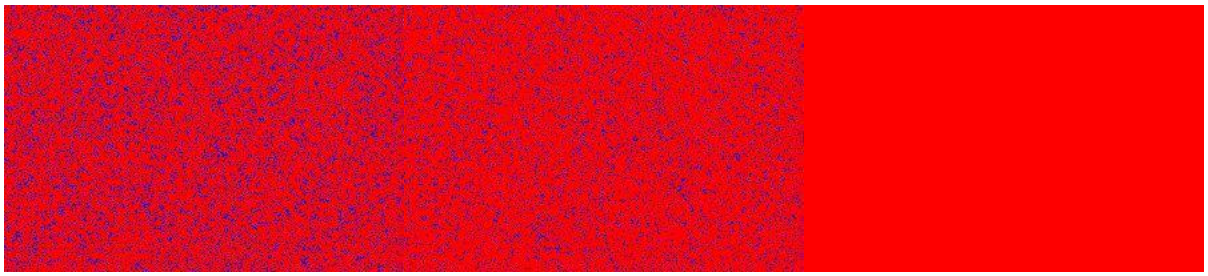
Red/Blue Dot Stimuli (80% Red)



57% 73% 76%



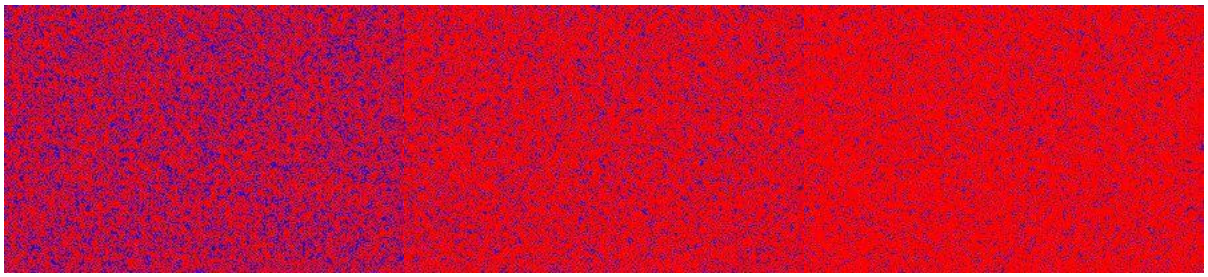
77% 80% 81%



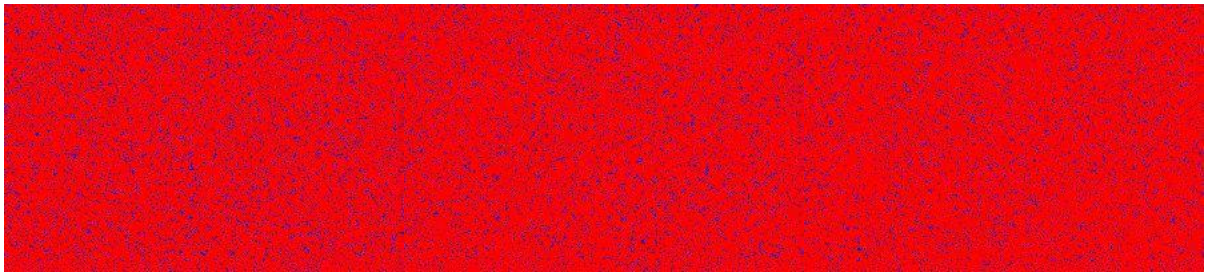
83% 89% 100%

Appendix O

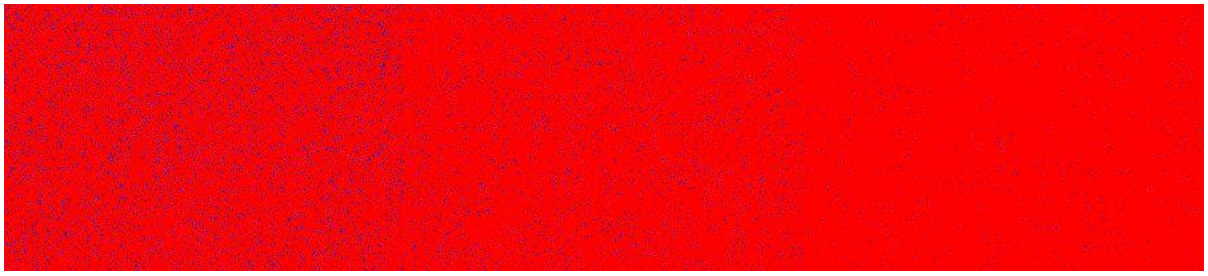
Red/Blue Dot Stimuli (90% Red)



72% 84% 89%



90% 91% 92%



94% 97% 99%