REMARKABLE U.S. ENGLISH LIKE ON A UNIVERSITY CAMPUS: NON-NATIVE USAGE, JUDGMENTS, AND ATTITUDES

By

Irina Zaykovskaya

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Second Language Studies—Doctor of Philosophy

2019

ABSTRACT

REMARKABLE U.S. ENGLISH LIKE ON A UNIVERSITY CAMPUS: NON-NATIVE USAGE, JUDGMENTS, AND ATTITUDES

By

Irina Zaykovskaya

Remarkable (D'Arcy, 2017) LIKE usages in English (i.e., as discourse marker/particle, quotative complementizer or approximator) comprise one of the "hot topics" of discourse-pragmatic research, but few studies have included L2 speakers (Diskin, 2017; Liao, 2009; Liu, 2016) and none of them looked at non-native attitudes or naturalness judgments. The current dissertation examined rLIKE as perceived, used and reflected upon by 26 NNSs – international students on a university campus with an average U.S. residence length greater than 40 months. Multiple data collection methods were used (interviews, syntactic judgment and matched-guise experiments); native speaker data collected via online questionnaire and from a local speech corpus were used for linguistic behavior, belief and attitude comparison. Results revealed that, despite a high level of within-group variation, NNSs behave native-like when using LIKE (token distribution across functions was similar for NSs and NNSs) and mostly native-like when judging naturalness of sentences containing LIKE in various syntactic positions. In addition to length of residence, beliefs about LIKE emerged as a factor influencing usage frequency in L2 speech: NNSs who perceived LIKE as specifically American were likely to use it often to signal belonging to the American English-speaking community. However, NNSs' level of stylistic awareness was low: most did not recognize LIKE as a stigmatized vernacular element that NSs prefer to avoid in formal situations. The attitude pattern (judgments about speakers using/not using LIKE) displayed by NNSs, however, was based on perceived social personae of speakers and native-like.

ACKNOWLEDGEMENTS

Eight years ago, on a cold November night in St. Petersburg, Russia, fresh from defending my first dissertation, I was feeling blissfully happy, because I was then sure that a life full of quiet library research was ahead. Little had I known then that there was another journey ahead of me, the journey that would involve moving across the ocean, going back to school, overcoming my fear of numbers, and finding out that conducting experiments and interviewing people would turn out to be even more exciting than spending nights at the library.

Along this journey, I met many wonderful people, but two of them (who are, incidentally, my dissertation co-advisors) stand out. Dr. Susan Gass, who was so welcoming when I first thought about applying to Michigan State University and sent her a cold email (ignorance was bliss, had I known then quite how legendary a researcher she is, I would have probably been too shy to do so!), and who agreed to be one of my advisors, even though my sociolinguistic proposal was not exactly within the area of her own interest, and who invested immeasurable amounts of time and thought into helping me make this dissertation happen. And Dr. Suzanne Evans Wagner, who made me fall in love at first class with variationist sociolinguistics and later became my advisor and mentor, and whose thorough and insightful feedback, as well as her warm smile and a shoulder to (occasionally) cry on, were invaluable. Thank you, Sue and Suzanne, for walking with me and making sure I did not stray from the path.

I also want to express my gratitude to Dr. Patti Spinner, a formal SLA researcher, and Dr. Peter De Costa, a non-variationist sociolinguist, for helping me develop my very first, pre-dissertation research projects (and making sure they were methodologically robust), and for their courageous agreement to be the members of my dissertation committee and dive into the depths

of discourse pragmatics with me. The multitude of perspectives is what makes my transdisciplinary soul rejoice!

Right next to my current committee, I want to acknowledge Dr. Valentina Chernyak, my advisor back in Russia, whose work ethic, encyclopedic knowledge, and kindness helped me become the person and researcher I am today. Also, I cannot be thankful enough to Drs. Aline Godfroid and Shawn Loewen, who somehow managed to teach me statistics in a way that I understood (which surely must have involved sorcery of some sort), to Dr. Debra Hardison, whose SPSS guidebook has been my Bible for the last few years, and to my amazingly smart fellow student, Xiaowan Zhang, who spent so much time helping me explore the best ways to approach my data.

This project would not have happened if Dr. Alexandra D'Arcy had not shared with me the entire manuscript of her then-unpublished but already incredible book "Discourse-Pragmatic Variation in Context: Eight Hundred Years of LIKE". It was truly an honor to study LIKE with a blessing from the LIKE Queen. I must also thank Drs. Kate Beeching, Chloé Diskin, and Devyani Sharma for sharing their unpublished works with me. It is difficult to find the right words to express what it means for a fledgling researcher to feel included into the academic community, so I'll just say that it means a lot.

Many other people offered me very real and much needed aid throughout my journey. Dr. Derek McNish and brilliant young actors Abby Byrne, Annie Courtney, Lee Cleaveland, Kayla Katona, Jenny Popovich, Marshall Ross, and Elise Jorgensen helped me prepare materials for my listening experiments. Scott Nelson, your help in the development and testing of the coding protocols was invaluable, and I would not have been able to transcribe those hours of recorded data without Jared Kaczor and several undergraduate transcribers whose work was funded

through MSU College of Arts and Letters Undergraduate Research Initiative (CAL-URI) grants in 2017-2018 to Dr. Suzanne Evans Wagner.

My deepest gratitude and appreciation go to my fellow members of SOSLAP, a student organization of Second Language Studies students at MSU, which is an incredible community that provided me with so many opportunities for professional growth and helped me navigate the treacherous path of graduate school. Another organization I must thank is SocioLab, which welcomed me as if I wasn't, technically, an outsider, and so warmly that every meeting filled me with inspiration and ensured a great mood throughout the day. Monica Nesbitt, your razor-sharp mind of a born linguist always left me in awe. Becca Senn, your infectious smile used to make bleak wintry Michigan days in the office brighter (and thanks for cat-sitting in times of need!). Delightful and mischievous Cara Feldsher, I found in you an office mate to talk about dinosaurs and weird things with!

For the first three years at MSU, I worked for the Russian program. The amount of support, understanding, and freedom to grow as a teacher I received throughout these years is indescribable. Drs. Jason Merrill, Shannon Spasova, and Zarema Kumakhova made my life as a graduate student juggling work and research as easy as it can possibly get.

Of course, a journey this long is unthinkable without friends and people who care about you just because. Suzanne Johnston, my first American friend and, quite possibly, a long-lost twin, with whom we discuss Universal Grammar, Harry Potter, and true crime books with equal passion. Jenn Brooke, now Majorana, my roommate of two years and the woman who got ordained just so I could get married. Hima Rawal, the gentlest and kindest person I have ever met and a great co-author with whom I published my first article. Karolina Achirri, the ball of energy who has never said "no" to a suggestion of coffee and talk. Dr. Paula Winke, who, while

not being my advisor, took it upon herself to make sure that I did not miss out of a fellowship opportunity (not to mention the inspiration her activism in the area of non-native speakers' rights provided). My friends Marina Krylova and Yulia Tchikhalova who proved that time difference and an ocean between us do not make our friendship any less close or meaningful and who were always there for me whenever I complained about all those new things I was experiencing in a foreign country. My family back in Russia, especially my grandma, who made enormous sacrifices to raise me and give me the best education possible. My cats, Gus and Scrat, who were literally by my side throughout the entire writing process, sometimes snoring, sometimes fighting, but always making me smile and melt a little.

I would also like to thank my entire cohort, most of whom are doctors by now, but I will always remember them as Susie – fellow cat lover – Kim, Zack – partner in academic crime – Miller, Dustin – the Thanksgiving Turkey Master – Crowther, Magda – brave soul who let me drive her car before I got a license – Tigchelaar, Jessica – we co-edited one hell of an MSU Working Papers issue together – Fox Brooker, Jie – OMG, she understands what island constraints are – Liu, and Jeff – father of Rosa, our cohort baby, - Maloney. Our classes together were pure delight.

Finally, there is my best friend, my husband, and my soul to acknowledge. Filipp, you helped me through so much, both physically and mentally, and you always made sure to remind me, to paraphrase our favorite J.K. Rowling, to turn on the light even in the darkest of times. Actually, you were that light, and this dissertation is at least half yours.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

aLIKE	Approximative Adverb LIKE
ANOVA	Analysis of Variance
AP	Adjective Phrase
CEFR	Common European Framework of Reference for Languages
CP	Complementizer Phrase
DCVE	Diachronic Corpus of Victoria English
DegP	Degree Phrase
DP	Determiner Phrase
EEG	Electroencephalogram
ESL	English as a Second Language
ESOL	English for Speakers of Other Languages
ICE	International Corpus of English
IELTS	International English Language Testing System
IHELP-MI	Influence of Higher Education on Local Phonology - Michigan
LINDSEI	Louvain International Database of Spoken English Interlanguage
LOR	Length of Residence
LVC	Language Variation and Change
mLIKE	Discourse Marker LIKE
MSU	Michigan State University
NNS	Non-Native Speakers
nP	Noun Phrase

NS	Native Speakers
pLIKE	Discourse Particle LIKE
PP	Prepositional Phrase
qLIKE	Quotative Complementizer LIKE
rLIKE	Remarkable LIKE
RM ANOVA	Repeated Measures Analysis of Variance
SD	Standard Deviation
SJ	Syntactic Judgment Experiment
SLA	Second Language Acquisition
TEA	Toronto English Archive
TOEFL	Test of English as a Foreign Language
TP	Tense Phrase
vP	Verb Phrase

CHAPTER ONE

INTRODUCTION

Second language research has begun to meaningfully grapple with acquisition of sociolinguistic variation and language change only fairly recently, even though the first widely known study on acquisition of a sociolinguistic variable by non-native speakers of English was conducted over forty years ago (Dickerson, 1975). The obvious obstacles preventing sociolinguistic research with non-native speakers from becoming mainstream are two- or even three-fold. One is the difficulty of distinguishing between linguistic and sociolinguistic competence (level of proficiency, on the one hand, and the nativelikeness of certain linguistic choices, on the other hand), which implies a related difficulty of developing a research design that would mitigate potential overlap between these types of competence. Another problem stems from the tension between the traditional reliance of sociolinguistic research on unmonitored speech and the desire to control as many potentially confounding factors as possible, coupled with a preference for experimental design, in much second language acquisition (SLA) research.

My dissertation will employ a research design that overcomes these issues by targeting highly proficient speakers (international students on a U.S. campus, fully enrolled in classes, for which a certain proficiency level must be achieved) and using a combination of experimental methods and interviews aimed at collecting unmonitored speech. I focus in this dissertation on non-native speakers' acquisition of 'remarkable' LIKE (term coined by D'Arcy (2017) and inspired by the fact that native speakers notice its presence in speech), a common element of the English vernacular that is known to be highly salient to native speakers and to be perceptually distinguished from, for example, the lexical verb (1a). 'Remarkable LIKE' is an umbrella term

used to refer to several different types of LIKE, the most important of which are illustrated by the examples below (1b-1e) and will be further discussed later.

(1) a. *I like dogs* (lexical verb)

(2) *b.* *I've had LIKE five cats* (approximator LIKE)

c. *And I was LIKE, "Are you kidding me?"* (quotative LIKE)

d. *LIKE, we went to the mall yesterday* (discourse marker LIKE)

e. *My sister and I are really LIKE close* (discourse particle LIKE)

The choice of remarkable LIKE, henceforward 'rLIKE', as my object of study is motivated by several factors. First, it is ubiquitous in native speech, which ensures its abundant presence in the input international students receive in their daily life on campus. As a socially salient[1] phenomenon (i.e., speakers are largely aware of the social information it carries), it is also extensively studied, so plenty of information is available with regard to the native patterns of its usage. At the same time, it is not taught to non-native speakers in instructed settings and = does not appear in textbooks and other educational materials, which allows me to assume that non-native speakers will be exposed to it only in an immersive setting (i.e., the setting of this study). This creates a unique research opportunity to compare native and non-native speech within the same linguistic environment. Finally, as a discourse-pragmatic (rather than a morphosyntactic or phonological) phenomenon, it is not inextricably embedded into the grammatical structure of English, which helps disentangle the patterns of sociolinguistic variation from grammatical accuracy of non-native speech.

In the following sections of this Introduction, I first outline some background on relevant transdisciplinary research in the fields of second language acquisition (SLA) and variationist

[1] In SLA research, salience may be understood differently; see Gass, Spinner, and Behney (2018) and Spinner, Behney, and Gass (2018) for review.

sociolinguistics, which is also known as the study of language variation and change (LVC). I describe theoretical and methodological issues that arise in this type of research and ways in which previous researchers have attempted to overcome them. I then present an overview of existing studies of discourse-pragmatic variation in non-native speech and introduce rLIKE. Finally, I briefly outline the structure of this dissertation.

SLA/LVC Connections and the Need for Transdisciplinarity

Variationist sociolinguistics studies sociolinguistic variables. A sociolinguistic variable, according to the most basic definition, is two or more ways of saying the same thing in a language, such as saying [kæt] vs. [kæʔ] for "cat", using "I'm going to sing" vs. "I'll sing" to express future temporal reference, etc. (Labov, 1972a). Each of these ways of saying the same thing, that is, a particular linguistic form, is called a variant, and a given form may function as a variant of multiple sociolinguistic variants (e.g., as an approximator, LIKE competes with *about* and *approximately*, while as a quotative, it competes with *say*, *be all*, etc.). Sociolinguistic variation, therefore, refers to the choices speakers make "when selecting the [linguistic] forms necessary to convey a message that is appropriate in a given context" (Geeslin & Long, 2014, p. 3)

Sociolinguistic competence, understood as "made up of two sets of rules: sociocultural rules of use and rules of discourse" (Canale & Swain, 1980, p. 30) has long been recognized as a crucial component of communicative competence, along with grammatical and strategic competences. However, it has been common practice in SLA to reinterpret sociolinguistic competence as primarily pragmatic and/or sociopragmatic competence. For example, Taguchi and Roever (2017), referring, among others, to Canale and Swain's (1980) model of communicative competence specifically, state that "models of communicative competence …

have situated pragmatic competence as an essential component of L2 ability. These models emphasize that communication missteps can occur from not understanding social conventions or rules of communication" (p. 7). As a result, L2 pragmatics and studies of intercultural competence have been at the forefront of the studies of communicative competence in general while sociolinguistic variation in the narrow sense, as a choice between specific linguistic forms in a given context, has largely remained beyond the scope of this type of research.

Despite decades of calls for more and deeper inter-disciplinary connections between the fields of SLA and variationist sociolinguistics, there has been little integration. Below are only a few selected quotations from papers and books published across the time span of over 40 years, each of which laments the lack of studies investigating sociolinguistic variation in non-native speech:

- "The variability model used in sociolinguistics has been used extensively in studies of sound change in the native language, but little has been done to apply this model to second language acquisition (Labov 1965, 1972)" (Dickerson, 1975, p. 402).
- "One of the most neglected areas of inquiry in the field of second language acquisition…is sociolinguistic variation" (Beebe, 1980, p. 433)
- "This volume ["Second Language Acquisition and Linguistic Variation"] represents an attempt to correct the relative neglect in SLA studies of the insights to be gained from the quantitative study of inter-language variation and to encourage interdisciplinary dialogue between researchers in SLA and in sociolinguistics" (Preston & Bayley, 1996, p. xiv).
- "Despite the apparent convergence of interest in understanding variability in language and in the speech of socially marginalized groups and individuals, until recently

variationist linguistics has had relatively little influence on SLA research" (Bayley, 2005, p. 2)

- "Since the beginning of research in Second Language Acquisition (SLA), we have learnt a lot about language learning, but there are aspects that have received less attention until recently. One of these is the acquisition of sociolinguistic competence" (Regan, Howard, & Lemée, 2009, p. 2)
- "Despite the abundance of research mentioned in Chapter One [of the book "Sociolinguistics and Second Language Acquisition: Learning to Use Language in Context"] that points to the importance of social factors in influencing second language acquisition and use, studies that focus on the connection between models of learner language, acquisition, and use, on the one hand, and the effects of various social factors, on the other hand, are quite scarce" (Geeslin & Long, 2014)
- "[Throughout the history of SLA,] the legacy of linguistics and psychology meant that most theories and insights remained strongly cognitive in orientation and generally ignored other research, such as Labov's (1970, 1972) in variationist sociolinguistics" (The Douglas Fir Group, 2016, p. 19)

But there are some promising signs. The actual number of variationist studies of non-native speakers has been slowly but steadily increasing over the years. The complexity of the process of acquiring sociolinguistic competence, as summarized by Meyerhoff and Schleef (2012), stems from the fact that "sounding 'native-like' requires (in no particular order): (1) matching native speakers' (NS) relative frequencies of key variants, and (2) matching the linguistic and (3) non-linguistic constraints on the variants in NS speech" (p. 399). The existing studies of this process "suggest that variation in interlanguage is not random but is highly

systematic and constrained by a range of linguistic and social factors", and "while these studies have shown that native variation is rarely replicated exactly, the patterns of variation used in the interlanguage of L2 learners do approximate those of native speakers" (Schleef, 2013, pp. 305-306) . Nonetheless, there is much to be learned about the mechanisms by which non-native speakers acquire the linguistic patterning and social constraints on variation in the target language. Advancing this knowledge, as stated in the previous section, will require SLA and sociolinguistic researchers to work together to overcome methodological problems.

Methodological Obstacles

In a landmark study of the acquisition of variable (*ing*) (i.e., the alternation of standard [ɪŋ] with non-standard [ɪn] in *running*, *ceiling*, etc.) by Vietnamese and Cambodian immigrants in Philadelphia, Adamson and Regan (1991) employed the previously made distinction between vertical and horizontal variation (Corder, 1981), or continuum (Young, 1988). According to the authors, "the study of the vertical continuum is the study of linguistic competence, and the study of the horizontal continuum is the study of sociolinguistic competence" (Adamson & Regan, 1991, p. 3). In that understanding, the progress along the vertical continuum is what the overwhelming majority of SLA researchers are concerned with. The overarching idea is that under favorable conditions (sufficient amount of input, plentiful opportunities for interaction, appropriate language instruction, etc.), most learners would produce more native-like forms over time.

However, the goal of separating the two types of continua (vertical and horizontal) for examination in SLA research continues to be difficult to attain. Firstly, the horizontal continuum may simply not be part of the acquisition process for participants under study. Foreign language learning that typically occurs in non-immersion instructed settings facilitates only the progress of

learners along the vertical continuum, with the notable exception of certain "poster" (stereotypical) differences (mostly lexical or spelling-related) between major varieties of the language studied, such as British vs. American English, or Metropolitan vs. Quebecois French. Only in the case of naturalistic learning, or instructed learning in an immersed environment (e.g., during study abroad), are learners simultaneously exposed to the vertical variation between native-like and non-nativelike forms (in extreme cases, to the point of complete communicative breakdown between native and non-native speakers) and to the horizontal, sociolinguistic variation (between different native speakers in a highly diverse setting, or between the variety learned prior to arrival and the local variety).

Secondly, L2 data on acquisition of variation may be compromised by confounds. For example, certain "standard" (native-like) variants may be articulatorily or otherwise difficult for most or some non-native speakers: An apparent failure of some L1 Chinese speakers of English to correctly supply regular past tense verb marking may be a result of pronunciation difficulties related to final consonant clusters (e.g. in *jumped* [dʒʌmpt]), rather than a signal of non-nativelike mental representation of English past tense. In their own study, Adamson and Regan (1991) resolved this issue by implementing a research design in which "the learners' native language supplies the prestige variant of a sociolinguistically sensitive form" (p. 3): L1 Vietnamese and Cambodian speakers they recruited had [ɪŋ], the prestige variant of the (*ing*) variable, in their native phonology, so that the appearance of nonprestigious [ɪn] in their English speech would unequivocally indicate the adoption of the local norm. This approach, however, severely limits research opportunities, and cannot thus be treated as a universal "gold standard".

Many typical SLA research methods also introduce confounds that make investigation of sociolinguistic competence difficult. Various experimental techniques (e.g., self-paced reading or

listening, eye-tracking, EEG or neuroimaging, such as fMRI, etc.) are used to investigate different aspects of linguistic (grammatical) competence of non-native speakers without collecting the production data that may potentially be contaminated by articulatory or other irrelevant (to the competence) difficulties. Studying sociolinguistic competence through speakers' perceptions and/or attitudes rather than their production, however, is complicated by the same issue – namely, the confounding of the vertical and horizontal continua. Thus, in a study of attitudes towards Irish English expressed by Polish immigrants, Diskin and Regan (2017) found that speakers with a higher level of overall proficiency in English tend to treat Irish English as a marked variety (as compared to the British English they learned in school, or American English they hear in the movies and TV shows) and express overtly positive or negative attitudes towards it. Speakers with a lower level of proficiency (and correspondingly shorter length of residence in Ireland), however, are more likely to be ambivalent in their attitudes and conflate their "general unfamiliarity with the language" with "unfamiliarity with a particular variety of that language" (p. 202). Although the authors targeted only "large-scale" attitudes towards a language variety as a whole rather than any particular variant of a linguistic variable, it seems justified to expect a similar issue of proficiency to affect sociolinguistic perceptions on the lower level, that is, on the level of concrete linguistic variables, as well. It is important to acknowledge that in Diskin and Regan's (2017) study, the proficiency was self-reported and generally correlated with the length of stay in Ireland: The speakers who lived in Ireland for over five years tended to assess their proficiency as advanced (C1, under Common European Framework of Reference for Languages), and speakers who lived in Ireland for only four years reported intermediate proficiency (B1-B2, under CEFR). The distinction between proficiency and length of stay in L2 country, however, may be especially sensitive for

sociolinguistic research, as only length of stay may be treated as an (imperfect) proxy to exposure to the local variety of the target language, while even high level of proficiency can be acquired outside of the target language environment.

Therefore, in addition to targeting both production and perceptions, any research aiming to assess the sociolinguistic competence of non-native speakers, as well as their attitudes towards specific language varieties and/or variants of particular sociolinguistic variables, needs to target highly proficient speakers to eliminate the low proficiency effect and ensure that the participants were exposed to a certain variety of language for a substantial amount of time. Two years of exposure seem to be a common choice among researchers (Drummond, 2011; Hellermann & Vergun, 2007; Schleef, 2013). International students on college campuses are an ideal population for such studies: in order to be enrolled, college students, unlike, for example, schoolchildren, need to demonstrate certain (usually high) baseline proficiency in the target language, and then they come to live in the target language environment for several years. Furthermore, while immigrants may, under certain circumstances, avoid using the target language at all, socializing strictly within their L1 environment and relying on family and community members to communicate with native speakers on their behalf, college students are exposed to and are forced to use their L2, at the very minimum, in classes and during obligatory events (orientation sessions, communication with advisors, etc.). This ensures the exposure to the native variation patterns in a naturalistic way. If the immediate linguistic environment (a college campus) is relatively homogenous with regard to the variant of the language spoken (e.g., the largest proportion of students on campus come from the same geographic area) and stable (e.g., no drastic changes in the proportion of native vs. non-native speakers on campus occurred within

several years), it should be considered conducive to non-native speakers' progress along the horizontal continuum.

Finally, a combination of various techniques of data collection (sociolinguistic interviews, ethnography, as well as experimental tasks) and data analysis (qualitative and quantitative methods) are necessary to effectively investigate the patterns of variation in non-native speech. Discussing the need for mixed-method approach in studies of sociolinguistic competence, Regan (2010) notes: "It is difficult to consider the significance of individual agency (for instance) in quantitative data alone, yet it is also difficult to satisfy the problems of generalisability and replicability if relying only on qualitative data. It is the combination of the two types of data that enables a nuanced analysis based on evidence, thus enabling interesting new insights into identity and the significance of L2 variation patterns" (p. 34). In fact, variationist sociolinguistics has a long-standing tradition of combining both methods of analysis, starting with Labov's (1963) study of Martha's Vineyard, which was essentially a quantitative analysis of variable pronunciation of two diphthongs, but employed qualitative observation of island life in order to arrive at the interpretation of the quantitative data. However, this type of research has been largely limited to studies with native speakers.

Research within the language variation and change paradigm began with the variation in phonology, which ensured a clear choice between functionally equivalent variants on the part of the speaker and allowed sociolinguists to focus on the social information behind speaker choices. Studying variation in grammar (morphology and syntax) has been more challenging due to the difficulties of circumscribing the so-called envelope of variation (all variants of the same variable) and ensuring the equivalence of different variants. In fact, Lavandera (1978) argued that, despite the indisputable value of research on variation in morphology and syntax, "it is

inadequate at the current state of sociolinguistic research to extend to other levels of analysis of variation the notion of sociolinguistic variable, originally developed on the basis of phonological data" (p. 171). Studying variation in discourse pragmatics has been even more problematic, because in most cases semantic or even functional equivalence of variants is very difficult to establish.

Discourse Pragmatics

Discourse-pragmatic features are a large class of linguistic elements, such as discourse markers, or particles (*well, you know, right,* etc.), general extenders (*that kind of thing, stuff like that, and* such, etc.) or adjectival intensifiers (*so, really, very,* etc.) are syntactically optional. Quotative verbs and expressions (*say, think, be like, be all,* etc.) are also considered discourse pragmatic variables. Discussing the scope of discourse pragmatics, Pichler (2013) states the following:

> "Discourse-pragmatic features constitute a formally heterogeneous category of syntactically optional elements which make little or no contribution to the truth-conditional meaning of their host units and – depending on their scope, linguistic co-text as well as sequential, situational and cognitive context – perform one or more of the following macro-functions: to express speaker stance; to guide utterance interpretation; and to structure discourse" (p. 4).

D'Arcy (2017) emphasizes that discourse-pragmatic features "are motivated by factors that exist in the mind of the speaker, as part of the online negotiation of meaning, which renders them difficult to reconstruct according to objective criteria" and that "these alternations sometimes behave like typical sociolinguistic variables, marked by form-function asymmetries, but sometimes they do not" (p. 1). This sometimes manifests in the lack of clear semantic or

pragmatic “meaning” or, in D’Arcy’s words, “equivalents”. Given this lack of objective criteria and the fact that most, if not all, discourse features are “categorically multifunctional” (Ajmer, 2002), it is not surprising that discourse pragmatics, a flourishing area of research, has only recently become fully integrated into the general variationist agenda. Indeed, the initial call for this integration that sparked “an unprecedented upsurge in quantitative research investigating patterns of variation and change in the use of conventionalized, polyfunctional linguistic items and constructions” (Pichler, 2016, p. 1) is widely attributed to Macaulay's (2002) chapter “Discourse variation” published in “The Handbook of Language Variation and Change” (Chambers, Trudgill, & Schilling-Estes, 2002).

Discourse Pragmatics in L2 Speech

The usage of discourse-pragmatic features in non-native speech began to attract researchers’ attention much earlier than discourse pragmatics was fully accepted as a subfield of variationist sociolinguistics. Approximately thirty years ago, multiple studies almost simultaneously declared a need to investigate how discourse marking contributed to the comprehensibility of the speech of international teaching assistants on American college campuses (A. Tyler, 1992; A. E. Tyler, Jefferies, & Davies, 1988; Williams, 1992) or uncover the general role of discourse markers “in establishing coherence in spoken [L2] English” (Hays, 1992, p. 24). Another line of early research was devoted to the perception of so-called academic discourse markers, or “lecturer’s OK” (Levin & Gray, 1983; replicated by Schleef, 2008), which were shown to facilitate understanding by both native and non-native listeners (Chaudron & Richards, 1986; Flowerdew & Tauroza, 1995).

Most studies investigating discourse markers in non-native speech quote Svartvik's (1980) who says that from a native speaker’s perspective, a non-native speaker who fails to

supply a *well* in a certain context (e.g., at the beginning of the reply to an uncomfortable question) may seem "dogmatic, impolite, boring, awkward to talk to" (p. 171, cited by Diskin (2017), p. 145). The non-native speaker would not be corrected and is, thus, unlikely to learn that the "form" she produced (zero marker instead of *well*) was, in fact, non-nativelike. Therefore, the acquisition of discourse-pragmatic features associated with a certain variety of the target language or a certain type of vernacular may serve as good indicators of the progress along the horizontal continuum, that is, the development of the sociolinguistic competence. Sankoff et al. (1997) suggested that "only L2 speakers with a high degree of contact with native speakers will master the use of discourse markers" (p. 193), which may also explain why many researchers choose to recruit students (college students in a study abroad setting or school-age immigrants) for the discourse-pragmatic acquisition studies, as educational settings guarantee regular exposure to the target language variety produced by native speakers from the same age group to which the non-native speakers belong.

While college students are overwhelmingly represented in second language acquisition research in general (Plonsky, 2015, 2017), most of the studies concerned with the acquisition of discourse pragmatics by non-native speakers tend to sample non-college populations, such as Anglophones in Quebec, Canada (Sankoff et al., 1997), Polish immigrants in urban and rural settings in Ireland (Nestor, 2013; Nestor, Chasaide, & Regan, 2012; Nestor & Regan, 2015) or Polish teenagers in Scotland (Truesdale & Meyerhoff, 2015), Polish and Chinese immigrants in Dublin (Diskin, 2013, 2017; Diskin & Regan, 2015), immigrants of various generations in the United States (Hellermann & Vergun, 2007; Lee, 1999).

Investigating the acquisition of variation (including discourse pragmatics) by students in immersive settings is, in fact, a fairly recent trend. It was started largely by Müller (2005), who

based her book on discourse markers on the Giessen-Long Beach Chaplin Corpus (GLBCC) which included recordings of non-native English speakers (mostly L1 German speakers) during their study abroad in Long Beach (California, USA). Davydova and Buchstaller (2015) studied the use of quotative markers by L1 German speakers of English represented in the Mannheim Corpus of German English (MaCGE), which contains recordings of German students "enrolled in degree programs with the highest uptake of study abroad programs" (p. 445). Beeching (2015) analyzed the use of discourse marker *well* by international students with various (but mainly Chinese) linguistic backgrounds at the University of the West of England. Fuller (2003) and Liu (2016) studied the use of discourse markers by graduate students (i.e., people who socialize as professionals rather than as students) at a U.S. university, while Liao (2009) researched teaching assistants (i.e., graduate students who are in the process of professional integration into the host community). In 2009, Regan et al. published a large collection of case studies on the acquisition of different variable features by L2 speakers of French along with several chapters establishing theoretical foundation for further research on the acquisition of sociolinguistic competence during study abroad.

Importantly, most studies listed in the two previous paragraphs have one thing in common: they all feature remarkable LIKE as one of the discourse elements under study (it applies to all studies on L2 English with the exception of Beeching's (2015) study which focused on a single discourse marker, *well*). The current study, however, will focus on rLIKE specifically and will investigate various aspects of its presence in non-native repertoire. In the next section, I will briefly introduce remarkable LIKE as a discourse-pragmatic phenomenon and the scope of research devoted to it as an element of native English speech. More detailed description of rLIKE functions and meanings, as well as attitudes towards it, and an extensive account of all studies of

rLIKE that involved non-native speakers will be presented in the Literature Review (Chapter Two).

Remarkable LIKE

LIKE is one of the most socially salient features of contemporary English language. Many – mostly pejorative – stereotypes about LIKE exist and are widely circulated in the media regardless of their correctness. In the foreword to her book "Discourse-Pragmatic Variation in Context: Eight Hundred Years of LIKE", Alexandra D'Arcy (2017) remembered that as a Ph.D. student choosing a topic for her dissertation, she was anxious at first when her supervisor suggested that she should look at LIKE. Among other things, she was scared "by the possibility of becoming "The LIKE Person". The overt policing of LIKE made it clear that this was a feature of vapid young women who simply injected LIKE in their talk when they couldn't pin down a thought" (p. xi). In addition to being salient, LIKE is highly frequent, to the point of being often described as ubiquitous, in native English speech, especially among adolescents and young adults (D'Arcy, 2017; Tagliamonte, 2016).

As a discourse phenomenon, LIKE began to attract the attention of sociolinguists in the mid-80s (Butters, 1982; Schourup, 1985; Underhill, 1988), and quickly became one of the "hot topics" of variationist research due to its apparent ubiquity, social saliency, and functional versatility. Thirty years later, Dinkin (2016) described the amount of research done on LIKE as "enormous"; the reference list of D'Arcy's "Eight hundred years of *like*" (2017) spreads over fifty-five pages. All possible aspects of LIKE are being researched, from perceptions of and attitudes towards it (Buchstaller, 2006; Dailey-O'Cain, 2000; Fox Tree, 2007; Hesson & Shellgren, 2015; Maddeaux & Dinkin, 2017) to its development in diachrony and grammaticalization patterns (Buchstaller, 2014; Meehan, 1991; Romaine & Lange, 1991;

Tagliamonte & D'Arcy, 2007) to phonetic realizations of LIKE as signals of style shifting (Drager, 2016). It is, undoubtedly, symbolic that, trying to build a new sociolinguistic methodology, (Dinkin, 2016) chose LIKE as a perfect case study "to illustrate the value of a variant-centered analysis" (p. 221).

The term "remarkable LIKE" was coined by D'Arcy (2005, 2007, 2017) in order to fulfill the need for the formal distinction between the socially salient type of LIKE frequently used by adolescents and young adults and the word "like" used as a verb or a comparative complementizer (*I LIKE dogs; I felt LIKE my life was falling apart*). In her doctoral dissertation, D'Arcy (2005) proposed an elegant way to untwine the tangle of LIKEs by dividing the possible uses of LIKE into two large groups based on the sociolinguistic status of each, which could be either "grammatical", that is, accounted for by prescriptive grammar of English (verb, noun, adjective, preposition, comparative complementizer, conjunction, suffix), or "vernacular". In her later works (e.g., D'Arcy, 2017), the terms were replaced with "unremarkable" (former "grammatical") and "remarkable" (former "vernacular"). Following D'Arcy, I adopt the term "remarkable LIKE" , although such terms as discourse LIKE, discourse-pragmatic LIKE, and other variants are frequently used in the literature (see overview in Diskin, 2017).

The main reason for my choice of "remarkable LIKE" over other available terms is that it comes with a fully-developed syntactic-functional typology based on the forms LIKE can take, as well as the roles each of those forms can play in a sentence and syntactic position(s) it can occupy. D'Arcy (2017) distinguished five major remarkable functions of LIKE. I illustrate them below with examples selected from my own data set of non-native speakers of English (Chapter Four):

(2) a. Quotative complementizer (*be like* preceding direct speech, thought, or mimetic

expression): *All I imagine is like I just do as police say and then he said, 'Oh yeah, someone told me about that, you look like a gang member.' So I'M LIKE, 'Oh great, thanks.'* (Ryan, m, 23, Vietnam)

b. Approximative adverb (used before numeric expressions): *There used to be a lot, LIKE a hundred or something, but right now really LIKE twenty people*. (Danni, f, 22, Thailand)

c. Discourse marker (clause-initial marker): *The community is good. LIKE there is Islamic Center right there, so I always go there*. (Aisya, f, 22, Malaysia)

d. Discourse particle (clause-medial marker): *And in a lot of LIKE foreign languages it's not LIKE the same LIKE sentence structure as English*. (Su, f, 23, Kenya)

e. Sentence adverb (clause-final marker, used primarily in Ireland, Northern Ireland, northeast of England, and Scotland): *You'd hit the mud on the bottom LIKE*. (example taken from D'Arcy, 2017)

I will present a more detailed account of forms and functions of rLIKE, as well as a review of literature devoted to its usage and factors affecting usage, beliefs about and attitudes towards rLIKE that speakers of English possess, in the next chapter of this dissertation. However, even from the brief overview presented above, it is clear that rLIKE is an interesting phenomenon to investigate with regard to non-native speakers, and in many ways, it is uniquely suited for this type of research. Thus,

(i) the ubiquity of rLIKE in native English speech ensures its steady supply in the input non-native speakers (NNSs) receive in an immersed setting,

(ii) the multifunctionality and intricate syntactic patterns of rLIKE usage provide an opportunity to investigate whether NNSs (who use rLIKE) behave in a truly native-

like fashion, rather than use it in a "pragmatically fossilized" (Romero Trillo, 2002) manner, and

(iii) social saliency opens a possibility of NNSs being exposed to explicit comments targeting rLIKE and frequent rLIKE users, which also creates an interesting line of research.

Dissertation Structure

This dissertation is organized in the following way. The Introduction is followed by a Literature Review chapter (Chapter Two), which presents a critical review of existing studies of rLIKE and reveals the gaps that exist in current scientific knowledge about the ways in which non-native speakers of English, especially international students on college campuses, use, perceive, and think about rLIKE. The following chapter, Methodological Overview, describes the context in which my study was conducted, introduces the participants of the study, and outlines the procedure. Due to the complicated methodological procedure and in order to facilitate the reading of this dissertation, the description of specific tasks, as well as treatment and analysis of collected data, is moved into Chapters Four and Five. For example, the description of a syntactic judgment experiment which was employed to analyze NNSs' sensitivity to various syntactic positions rLIKE may occupy, can be found in Chapter Four ("Remarkable LIKE in native and non-native speech: Usage and Syntactic Judgments"), while the matched-guise experiment, which was used to elicit attitudes towards rLIKE, will be described in Chapter Five ("Native and non-native beliefs about and attitudes towards remarkable LIKE"). The findings of the study and the discussion of these findings are presented in Chapters Four and Five. Conclusion contains a brief summary of the study results, as well as limitations of the study, pedagogical implications, and directions for future research. The

dissertation concludes with References and Appendices in which the instruments materials used in the study (e.g., interview protocols, instructions given to participants during experiments, etc.) are presented.

CHAPTER TWO

LITERATURE REVIEW

In this chapter, I will present a review of existing studies devoted to remarkable LIKE (which is frequently referred to as "discourse LIKE", "discourse-pragmatic LIKE", or "discourse marker LIKE" in the literature, even if researchers include quotative LIKE in the study). As the primary focus of this dissertation is on the ways in which non-native speakers of English interact with LIKE (use it, perceive it in the speech of others, or think about it), this literature review also highlights the studies that included non-native-speaking participants.

This chapter consists of two major parts, each of which contributes a different perspective on rLIKE. The first part, "Remarkable LIKE as a discourse-pragmatic feature", presents the linguistic view on rLIKE. This view is based on the fact that "the different functions of like participate in different variable systems" (e.g., as a quotative, "be LIKE" competes with other quotative verbs and expressions, such as "say" or "be all", but as a discourse marker preceding a clause, it does not), which "means that standard variationist methodology requires treating them separately" (Dinkin, 2016, p. 228). Following this logic, I discuss the existing approaches towards how the functions of rLIKE can be delineated and what methodological issues stem from the difference between these approaches. I then present a detailed summary of the approach I take in this study (which is the syntactic-functional approach by D'Arcy, 2017), and finally review the existing research on rLIKE usage with regard to its different functions.

The second part, "Remarkable LIKE as a carrier of social meaning", essentially presents the public view of rLIKE, because the studies reviewed in that part are not concerned with the differences between rLIKE functions, but rather reflect the perception of rLIKE. As Dinkin (2016) rightfully noted, "although the vernacular functions of *like* belong to different variable

contexts and have different covariants, general commentary on *like* by non-linguists indicates that, in overt evaluation, the different vernacular *likes* are not distinguished from each other, and share sociolinguistic evaluation" (p. 230). As such, studies in this section of the chapter are those that took "the people's view" in that they acknowledged the sociolinguistic reality of remarkable LIKE (as opposed to the unremarkable LIKE) as a single entity, perceived as such by the people who participated in those studies – notwithstanding the fact that they were conducted by trained sociolinguists and employed methodologically sound and robust research designs.

Remarkable LIKE as a Discourse-Pragmatic Feature

As I already wrote in the Introduction, remarkable LIKE was recognized as a discourse-pragmatic feature of the English language in the 1980s and since then has been researched extensively with regard to the native patterns of its usage (see bibliography in D'Arcy, 2017).

The study of non-native usage of rLIKE is a much more recent trend pioneered by Lee (1999) in his doctoral dissertation, in which he investigated the usage of three discourse markers (*like, I mean,* and *you know*) by Korean immigrants of three generations in the U.S. Soon after, Fuller (2003a) compared native and non-native usage of discourse markers *well, oh, y'know, like, I mean*. The first widely known and cited study of L2 usage of rLIKE, however, was Müller's (2005) book on the discourse markers (*so, well, you know*, and *like*) using the Giessen-Long Beach Chaplin Corpus which included the recordings of non-native English speakers (mostly L1 German speakers) during study abroad in Long Beach (California, USA). Only a few papers investigating non-native usage of rLIKE have been published since then, many of them featuring not only rLIKE but other discourse markers of quotative verbs as well. Most of them were concerned with immigrant language learners, such as Polish teenagers in Scotland (Truesdale & Meyerhoff, 2015), beginning adult English learners in the U.S. (Hellermann &

Vergun, 2007), Polish and Chinese immigrants in Ireland (Diskin, 2013, 2017; Diskin & Regan, 2015; Nestor, 2013; Nestor, Chasaide, & Regan, 2012; Nestor & Regan, 2015). Two looked at discourse marker use by graduate students on U.S. college campuses (Liao, 2009; Liu, 2016) and one analyzed the usage of quotatives (including *be like*) by German college students (Davydova & Buchstaller, 2015). The latter was one of the only two studies that targeted English speakers outside a primarily English-speaking country, the other being Algouzi's (2015) dissertation investigating the usage of discourse markers by Saudi learners of English. While Siemund, Maier, and Schweinberger (2009) used the corpora of Indian, Philippine, and East African Englishes (all countries where English is not the primary language as well), these corpora inevitably included not only L2 English speakers but also early bilinguals who grew up speaking both their local language and English, which distinguishes this study of rLIKE from other works with non-native-speaking participants. None of the studies I am aware of featured undergraduate students on an English-speaking college campus as rLIKE users. Thus, the present study described in this dissertation makes use of a novel type of non-native participant sample.

Due to major differences in participant demographics, sample size and methods of data collection, coding and analysis, the findings of the existing studies (which I will summarize later in this part of the chapter) cannot be easily compared, so only one claim can be made with absolute certainty: NNSs, at least those living or having lived in the target language environment, do use rLIKE in their speech. There is also evidence of NNSs being sensitive to locally-specific patterns of syntactic placement of rLIKE (Nestor, 2013; Nestor & Regan, 2015).

Methodological Issues of rLIKE Research

While published works investigating LIKE are abundant, there is remarkably little consistency in the literature with regard to the ways the usage of LIKE is analyzed. All

researchers first encounter the problem of distinguishing between various LIKEs that appear in the data, then the problem of coding the actual tokens of LIKE found in the data in the absence of any existing guidelines that are uniformly agreed upon, and, finally, the problem of interpreting calculated frequencies and established constraints without many comparable reference points (in terms of the types of LIKE taken into account and coding procedures which may not be fully and/or clearly reported). Naturally, this makes general findings of any such study (e.g., whether women use LIKE more frequently than men, what social, stylistic and linguistic constraints apply to the usage of LIKE) less comparable across studies, because the findings may be based on different data, depending on what types of LIKE were included or not included in the analysis.

Whenever the usage of LIKE by non-native speakers is concerned, an additional layer of problems is added on top, such as establishing the English proficiency of the participants and justifying selected level(s) or the decision to include all levels, controlling participants' exposure to native speech, usually via a proxy, such as length of residence or stay, or by relying on self-reports, and developing new or adapting existing coding procedures specifically for grammatically non-nativelike speech.

Even though only D'Arcy (2005, 2017) and her direct followers in the area of LIKE research (e.g., Maddeaux & Dinkin (2017)) openly and consistently employ the distinction between unremarkable ("grammatical") and remarkable ("vernacular") LIKE, most other researchers seem to imply this distinction as well. There is an overarching assumption that the difference is evident and needs not be further discussed, so researchers often use other terms (e.g., "discourse LIKE", or "discourse marker LIKE") as umbrella terms for all or most types of LIKE that D'Arcy would recognize as remarkable, often including even grammatically distinct

quotative LIKE under such terms. For example, Truesdale and Meyerhoff (2015) include quotative LIKEs into their study of discourse marker LIKE, justifying their decision by stating that "it seems fairly clear that the use of *like* as a quotative verb was made possible by the simulative or approximative functions of *like* elsewhere in the grammar" (p. 9) and suggesting that this inclusion may enable them "to consider the extent to which the enrichment of the variable in the speech of L2 users is the same or different from the enrichment of the functions of *like* over time" (p. 9). Nestor and Regan (2015), on the other hand, do not even mention the possibility of including quotatives when explaining their understanding of "discourse LIKE" (which, in their study, includes clause-initial, clause-medial, and clause-final markers). Studies that focus on quotative LIKE (e.g., Davydova & Buchstaller, 2015) typically do not calculate independent frequencies (i.e., per 1000 or 10,000 words) but compare the proportion of LIKE tokens used to the proportions of other quotative verbs or expressions. Many other studies of rLIKE also calculate proportions rather than frequency rates. For example, D'Arcy (2017) calculated the proportions of syntactic contexts (clauses or phrases) headed by rLIKE among all contexts of the same type. In some studies, other ways of reporting frequencies may be involved. For instance, Lee, (1999) divided the total number of words uttered by a group of participants by the number of discourse marker tokens; Hellermann and Vergun (2007) calculated the rate of using discourse markers per turn.

The choice of methodological approach itself (both with regard to data collection and data analysis) inevitably creates space for criticism with regard to how accurately the resulting statistics represent the "real-life" dynamics of LIKE usage. Truesdale and Meyerhoff (2015) summarized the problem of researching structurally optional elements such as discourse markers as "the problem with not having a clear envelope of variation" and proceeded to ask themselves

and the readers: “What is the denominator going to be when you can’t specify all and every place a speaker might choose to express politeness or intersubjective alignment?” (p. 11). Discussing D’Arcy's (2005) decision to apply the principle of accountability to her data and search for LIKE among the randomly and selectively sampled syntactic contexts, Truesdale and Meyerhoff (2015) further noted that while this approach allows “to control the denominator in a way that analyses of variation taken from naturally occurring conversations usually cannot”, “the cost is at the expense of full accountability to the numerator” (p. 11). In other words, no perfect solution to this problem exists.

Trying to ensure full accountability to the numerator, however, carries the intrinsic risk of task effects. For example, using a specific task rather than a generic sociolinguistic interview as a means of speech elicitation (e.g., narrating a movie, as in Müller (2005)) may promote the production of more discourse markers or quotative LIKEs than could be expected during a sociolinguistic interview. Multiple other factors, such as the level of formality set by the interviewer and the rapport established between the interviewer and the participants, may affect the results. When qualitative analysis of the content is conducted, it is possible to rethink the role of the interviewer and analyze his or her contribution to the interview dynamics as well, by treating the entire event not as a mere research tool but as a social practice (Talmy, 2010; Talmy & Richards, 2011). However, when the goal is as specific as eliciting certain elements of the vernacular and analyzing the linguistic and extralinguistic circumstances of their occurrence, this approach may not be appropriate as it does not allow for direct interspeaker comparability. Thus, certain ways of operationalizing the interviewer factor must be applied. For example, Fuller (2003) found that, contrary to her expectations, “*like* rates for the speakers were actually higher in their interviews than in the casual conversations they recorded with intimates”, and speculated

that LIKE could be especially useful in an interview setting when volumes of new information must be presented to the unfamiliar interlocutor.

Finally, a researcher usually chooses between two major approaches to establishing the differences between types of LIKE. To complicate things, both of them are often referred to as functional. The chronologically first approach originated in the works by Schourup (1985) and Underhill (1988) and relies mainly on pragmatics and semantics. For example, Romaine and Lange (1991) suggested recognizing two major functions of discourse LIKE: the "interpersonal", or focusing, function, and the quotative function. As follows from the terminology, the semantic (discursive) rather than syntactic function is used as a basis for categorization (e.g., non-contrastive focus). Levey (2006), in his analysis of the use of LIKE by preadolescents in a London school, distinguished such functions of discourse marker LIKE as exemplification, metalinguistic focus, and discourse connectivity, which also suggests semantic, or pragmatic, approach. Multiple studies investigating the usage of LIKE by both native and non-native speakers adopted this approach and contributed to the ever-growing list of functions, adding such functions as hesitation, exemplification, hedging, etc. (Fox Tree, 2010; Fuller, 2003; Hellermann & Vergun, 2007; Liu, 2016; Müller, 2005; Romaine & Lange, 1991; Truesdale & Meyerhoff, 2015; see summary in Diskin, 2017). Attempts were made to establish the "core" meaning that would account for the meaning all types of rLIKE: for example, Andersen (1998) suggested that LIKE "helps the hearer to consider (a certain element of) the proposition expressed as loosely used" (p. 166), while Beeching (2016) stated that the "overarching core function" of LIKE is "to flag approximation and hedge discourse" (p. 127).

Another approach, syntactic (or structural), is motivated by the fact that discourse markers "often display a startling degree of structural promiscuity" (Truesdale & Meyerhoff,

2015, p. 10) and focuses on the positional distribution of LIKE. This approach offers clean-cut coding criteria that do not involve subjective judgment on the part of the researcher and is thus well-suited for analyzing data which may lack context for establishing the pragmatic function of each token of LIKE (e.g., corpus data). This approach is most popular among the researchers of English varieties outside of North America because it allows for comparison of usage patterns between and within these varieties without relying on pragmatics certain aspects of which may be specific for each variety. Thus, Siemund, Maier, and Schweinberger (2009) compared the occurrence of LIKE in clause-initial, clause-medial and clause-final positions of spoken sections of ICE (International Corpus of English) Ireland, ICE India, ICE East Africa, and ICE Philippines and found that clause-final LIKE expectedly accounted for over 39% of all LIKE tokens in Irish English, and, quite unexpectedly, almost 20% of Indian English LIKEs. The authors used this unexpected finding to support their argument for the necessity of considering structural and functional parameters when analyzing areal distributions of non-standard features rather than merely establish their presence or absence in a given language variety.

More frequently, the analysis of the syntactic position of LIKE is used to investigate social factors influencing its usage. For example, various studies of Irish English showed different trends among older, male and rural and younger urban speakers, with the former favoring clause-final LIKE and the latter preferring LIKE in clause-medial positions (Amador-Moreno, 2012; Corrigan, 2015), and with non-native speakers displaying similar distributions (Nestor et al., 2012; Nestor & Regan, 2015). Some studies combine the pragmatic and syntactic approaches, usually by analyzing the pragmatic functions of LIKE while also distinguishing between clause-medial and clause-marginal positions (e.g., (Diskin, 2013, 2017; Levey, 2006; Nestor & Regan, 2015).

Among the various studies analyzing the syntactic position of LIKE, alone stands the functional approach proposed by D'Arcy (2005, 2017). It is unique not only because of its promise to apply Labov's principle of accountability to the study of LIKE, but also because of the clearly stated distinction between unremarkable and remarkable LIKE and the exhaustive list of syntactic contexts in which rLIKE may appear. While the already mentioned distinction between remarkable and unremarkable LIKE is the important first step of the analysis of how people use and perceive LIKE, another critical step is the morphosyntactic analysis of what forms LIKE can take, what role each of those forms can play in a sentence, and which syntactic position(s) it can occupy. Within this approach, a function is recognized as a type of LIKE that D'Arcy recognized through that analysis, and syntactic contexts in which each function of LIKE can be manifested, as well as contexts in which LIKE is prohibited or constrained, are established.

Remarkable LIKE: Forms and Functions

Although I presented the five major functions of rLIKE proposed by D'Arcy (2017) in the Introduction, it is important to present a detailed version of D'Arcy's syntactic-functional typology and introduce some abbreviated terms that will be useful for the analysis of usage and perception of rLIKE by syntactic position, which will be presented in Chapter Four of this dissertation.

As already stated in the Introduction, D'Arcy (2017) distinguished five remarkable functions of rLIKE (examples 3a-3e taken from D'Arcy (2007, 2017)):

(3) a. **Quotative complementizer (qLIKE)**: And we WERE LIKE, "Yeah but you get

to sleep like three-quarters of your life." He WAS LIKE, "That's an upside." [2/f/12][2]

b. **Approximative adverb (aLIKE)**: It could have taken you all day to go LIKE thirty miles. [N/f/76]

c. **Discourse marker (mLIKE)**: Nobody said a word. LIKE my first experience with death was this Italian family. [N/f/82]

d. **Discourse particle (pLIKE)**: Well you just cut out LIKE a girl figure and a boy figure and then you'd cut out LIKE a dress or a skirt or a coat, and like you'd color it. [N/f/75]

e. **Sentence adverb** (largely restricted to certain dialects of English spoken mainly in Ireland, Northern Ireland, northeast of England, and Scotland): You'd hit the mud on the bottom LIKE. (TEA/62m/1941)[3]

While quotative complementizer LIKE (2a) had long been recognized as a specific entity in sociolinguistics (Butters, 1982; Schourup, 1985), approximative adverb LIKE (2b) was not much discussed, and both clause-initial (2c) and clause-medial (2d) LIKEs were treated as the same phenomenon. These latter two cases were usually collectively labeled—along with such items as *well, right, oh, you know, I mean*—'discourse markers', 'pragmatic markers', or 'discourse-pragmatic markers', among other names. Indeed, Diskin (2013) lists eleven different terms she found in the literature (p. 69).

D'Arcy (2017) was the first researcher who proposed a formal distinction between discourse marker (2c) and discourse particle (2d). This distinction is largely based on differences

[2] "Parenthetical information following examples marks the subcorpus from which the datum was extracted, followed by speaker's sex and age" (D'Arcy, 2007, p. 413).
[3] This example was taken from D'Arcy (2017). TEA stands for "Toronto English Archive", and the remainder of the code indicates the speaker's age and gender at the time of data collection, followed by year of birth.

in syntactic distribution (the discourse marker occupies clause-initial position only, while the discourse particle may occur in a number of clause-medial positions) and function in discourse (mainly textual/structural for the discourse marker and interpersonal for the discourse particle), as well as diachronic development, as discussed in the next section.

Discourse Marker, mLIKE. According to the trajectory hypothesized by D'Arcy from her study of ten corpora of historical English and ten corpora of contemporary English, at the end of the eighteenth century the clause-final discourse marker (essentially, a grammaticalized sentence adverb) changed position and appeared "on the left periphery of matrix clauses, where it was favored in sentence initial position and introduced the main proposition" (D'Arcy, 2017, p. 157). From matrix CP (4a), it subsequently (at the beginning of the twentieth century) started to precede subordinate CPs (4b), and then, a little more than half a century later, appeared in the new adjunction site, the left edge of subordinate TP (4c), where it "intervenes between the material that is hosted in CP (*because*, *that*, etc.) and the subject of the clause" (p. 122). The examples below are all taken from D'Arcy (2017, pp. 120-122).

(4) a. LIKE we sort-of lost touch for half of high school, probably 'cause I went to Upper Canada College. (TEA/24m/1979)

b. So I get it all done [LIKE when I get home]. (TEA/17f/1986)

c. It's weird [because LIKE you didn't really fit in the black group]. (TEA/21f/1982)

The uncovering of such diachronic trajectories / hierarchies as matrix CP > subordinate CP > TP are not only invaluable for historical linguistics, but they have immediate implications for contemporary sociolinguistic research, because "once LIKE spreads to a new projection, that particular site is established for successive generations" (p. 159). This means that the youngest

speakers have the full range of adjunction sites at their disposal, while older speakers are likely to have a narrower range. That is, "if speakers exhibit a limited use of the marker [or a particle], they do not use it in one of the more recent adjunction sites" (p. 161). In short, the relationships between syntactic host sites for LIKE are implicational: A speaker's lack of use of any position implies lack of use of all positions that are diachronically newer hosts for LIKE. This can be illustrated by data from individual speakers taken from the Toronto English Archive (Tagliamonte & D'Arcy, 2007) represented in Figure 2.1. Commenting on this figure, D'Arcy notes that this implicational hierarchy "reflects the broader, community norm: If the TP context is used by a speaker, then both the CP subordinate and CP matrix contexts are attested in their vernacular practice as well. The reverse does not obtain" (p. 162). It is also notable that rLIKE in the TP context was not observed for speakers older than 30 at the time of data collection (2002 – 2006).

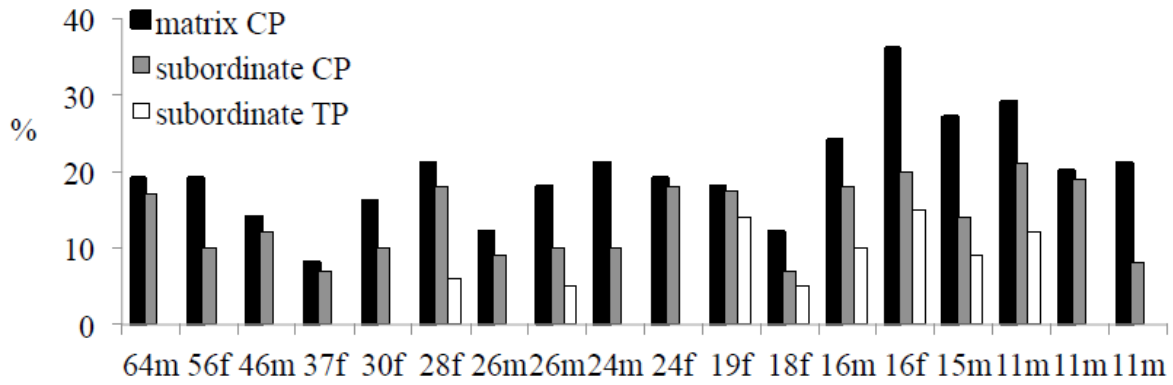

Figure 2.1
Distribution of marker like*: evidence from the individuals*. The figure taken from D'Arcy (2017, p. 162). The numbers stand for speaker age, the letters indicate gender.

The distributional data show that the oldest (and, hence, the most firmly established) adjunction sites are also ones that most frequently host LIKE in contemporary speech. In the TEA, which contains audio-recorded sociolinguistic interviews with 99 speakers born between

1916 and 1992, mLIKE appeared in 14.2% of all CPs, both matrix and subordinate, but only in 7.9% of subordinate TP clauses. Naturally, the frequencies of mLIKE in any particular adjunction site, including subordinate TP context, may be higher or lower in other varieties of English spoken outside of the Toronto area, but, given the historic (developmental, in (D'Arcy's (2017) terms) reasons behind this particular distributional pattern, the overall pattern is likely to hold true for other varieties as well. Therefore, the input that non-native speakers may be exposed to in the native-speaking environment in the United States is expected to also adhere to the distributional pattern in Toronto English. The results of a regular in-class experiment (native speaker naturalness judgments of LIKE in seven different syntactic positions) conducted by Wagner with U.S. English-speaking undergraduate students support this claim (personal communication, 08/22/2017).

Discourse Particle, pLIKE. The discourse particle, a result of further development of the discourse marker into other areas of the syntactic structure, may currently occupy many (though not all) clause-medial positions. Using data from the Diachronic Corpus of Victoria English (DCVE; D'Arcy, 2011-2014), D'Arcy showed that the left periphery of the DP (4a) and the vP (4a) are the oldest contexts, dating back to c. 1865. From the DP, the particle quickly spread to DegP (4b), then AP (4c), and later to nP (4d), that is., between a determiner and a noun.

4. a) I almost felt like I was cheated because I just like know how I'd act. (DCVE/90f/1865)

   b) I haven't seen LIKE a huge difference. (TEA/45m/1958)

   c) Like I love her but she's LIKE dumb. (TEA/18f/1985)

   d) I get really LIKE flabbergasted. (TEA/24f/1979)

   e) Like we were supposed to rememorize some LIKE parts. (TEA/11m/1991)

In her analysis of syntactic contexts in which pLIKE may appear, D'Arcy distinguishes three major domains based on the type(s) of phrase that host such contexts: nominal (DP, nP), adjectival (AP, DegP), and verbal domains (vP). In the following subsections, I will summarize the findings related to the nominal and verbal domains as I restricted my own research on the usage and judgments of pLIKE to these; the detailed account of the adjectival domain can be found in D'Arcy (2017, pp. 139-143).

Nominal Domain. Within the nominal domain, pLIKE accounted for 9.9% of all DPs produced by speakers born after 1925 (older speakers used pLIKE only in a negligible number of DPs). However, not all DPs are equally likely to receive pLIKE, as pLIKE "is significantly more frequent with arguments than with complements, suggesting that *like* may be more grammaticalized in the former than in the latter" (D'Arcy, 2017, p. 133). In DPs functioning as arguments (i.e., as a direct object), pLIKE appeared at a rate of 12.9%, while only 5.5% of DPs functioning as complements of prepositional phrases (PPs) included pLIKE. The least frequent adjunction site for pLIKE within the nominal domain, however, is the nP context, in which pLIKE can be inserted between a noun and a modifier. This type of insertion was observed in just 21 out of over 4000 DPs, which allowed D'Arcy to describe the nP context as a novel one. Of course, as over 15 years have passed since the data collection for the corpora D'Arcy used, it can be assumed that pLIKE in the nP context has become more frequent. In a recent perception study, Maddeaux and Dinkin (2017) found that the nP context was more salient than all other functions of LIKE, both remarkable and unremarkable, for young New Yorkers who participated in their experiment.

Verbal Domain. Within the verbal domain, pLIKE "categorically occurs to the immediate left of the lexical verb. When functional morphemes such as modal verbs, auxiliary verbs, and

infinitival *to* are present, *like* appears between these and the main verb" (p. 144). A constraint applies: pLIKE does not occur before finite inflected *be* and only occurs after the non-finite *be*. However, D'Arcy acknowledged that sporadic appearances of pLIKE before finite inflected *be* were observed in her data, which implies that the constraint is not, in fact, categorical, and further intrusion of pLIKE into the syntactic structure may be ongoing. Overall, pLIKE appeared in 7.6% of the vPs in TEA, which made it the third most frequent adjunction site for rLIKE after pre-CP mLIKE and pre-DP pLIKE.

Constraints on rLIKE Usage. As pLIKE continues to explore new adjunction sites (contexts) and yet at the same time, very little research focuses on syntactic positions of rLIKE, it is difficult to claim that any known constraints on rLIKE usage are categorical. However, it is safe to assume that rLIKE appears in certain contexts with much lower frequency than in other contexts. Furthermore, the differences in frequency may be assumed to exist between rLIKE in the older, well-established contexts (e.g., aLIKE before numbers and quantities, pre-CP mLIKE, pre-DP pLIKE or pLIKE on the left periphery of the lexical verb) and in the novel contexts, such as pre-nP and pre-subordinate TP. Finally, regional/dialectal differences may be expected: For example, sentence adverb (clause-final LIKE), while well-established and thus relatively frequent in Irish English, some dialects of British English, and Indian English (Siemund et al., 2009), may be virtually absent in American, Canadian or other varieties of English .

Factors Affecting rLIKE Usage

Native Speakers and rLIKE Usage. Schweinberger (2011) analyzed the International Corpus of English (ICE) data from eight regional varieties of English (U.S., British, Canadian, Irish, New Zealand, Jamaican, Indian, and Filipino) and found considerable differences in both the overall frequencies and syntactic positions in which discourse marker LIKE (quotatives

excluded) may appear. For example, the calculated frequency of LIKE in Canadian English was 4.38 instances per 1,000 words, while in British English it was only 0.49. At the same time, similar rates across different varieties did not necessarily imply the similarity of patterns: Thus, "instances of LIKE in CanE are almost exclusively confined to clause-initial and clause-medial uses of LIKE, while IrE owes its high overall rate to a substantial amount of the so-called traditional clause-final forms" (p. 380). Crucially, Schweinberger found that the distribution of discourse marker LIKE in different regional varieties of English was subject to different social and/or linguistic factors. Numerous studies of other kinds of remarkable LIKE, notably quotative LIKE, have also found considerable differences across regional English varieties in its social and/or linguistic patterning (see Durham et al. (2012) for British English; Singler (2001) for American English; Tagliamonte & D'Arcy (2004) for Canadian English; Winter (2002) for Australian English).

Besides region, the three other major social factors that are typically considered in sociolinguistic research are age, gender, style, and socioeconomic status. Of these, socioeconomic status has received curiously little attention in studies of rLIKE, although this may be due to the typical focus on adolescent and pre-adolescent speakers, whose status can often not be determined with reliability[4]. It will not be considered further here.

Age. Age, indeed, is a well-attested factor in the studies of rLIKE, given its rapid incursion into English vernacular use in the last century. According to Schweinberger (2011), "LIKE, although not limited to younger speakers, is significantly less likely to be used by older

[4] For example, although Schweinberger (2011) found that "substantial social stratification is observable in locales where LIKE is well established" (p. 396), because he used an independent corpus, only information about speakers' occupation was available to Schweinberger, so it was the only index of socioeconomic status that he used, and for the same reason he removed students, that is., a large portion of younger speakers, from the analysis (because their parents' occupation was not known).

speakers" (p. 386) across all varieties of English, a finding that applies even to the usage of clause-final LIKE in Irish English which is commonly associated with older and/or rural males. D'Arcy (2017) underscores that, despite being more common in the speech of younger people, rLIKE "is shared by the whole of the speech community and their forbearers, with most of its adjunction sites available to speakers of all ages, even if in some cases a specific context of use is rare" (D'Arcy, 2017, p. 165). The difference between older and younger speakers, as D'Arcy argues, is quantitative rather than qualitative.

Gender. Gender, on the other hand, is more nuanced and, as multiple studies show, may operate differently across different functions and syntactic positions of rLIKE and across different varieties of English, all while being intertwined with age, place of living (urban vs. rural), and style. Decades of sociolinguistic research have shown that women lead language changes, with parity across the sexes being achieved (if at all) only once the change is diachronically well-established (Labov, 2001a). Each of the many studies of LIKE has necessarily captured gender distributions associated with a specific point in LIKE's diachronic trajectory in a specific speech community, and even within a specific syntactic position or function. Thus, the LIKE literature appears at first blush to contain many contradictory findings regarding gender, but these fall out from the fact of LIKE's very rapid incrementation and diffusion in just a few decades.

In American English, for example, "none of the uses of LIKE are correlated with gender" (Schweinberger, 2011, p. 391), but in Canadian English, "the use of clause-initial LIKE … is strongly correlated with females, while other variants are not notably associated with gender" (p. 392). D'Arcy (2017) found a similar trend with regard to the clause-initial marker, but only after she removed speakers born in the 1950s and 1960s (middle-aged at the time of the study) from

her data and left only younger and older speakers in the sample. She found the difference to be "most prominent among the groups who are advancing the marker at the higher rate" (p. 169), that is, speakers born after 1980. The overall distribution of clause-initial discourse marker LIKE showed no significant difference between gender in her Canadian English data. Levey's (2006) research on the use of rLIKE by English preadolescents shows that gender differences may appear as early as at the age of 7-8 and become robust by the age of 10-11. In his data collected in London, 17% of all rLIKEs tokens produced by younger (7-8 y.o.) girls were the tokens of clause-initial LIKE (i.e., the discourse marker), while for younger boys that proportion was slightly lower (11%). Older (10-11 y.o.) girls, however, noticeably increased their usage of clause-initial LIKE (32% of all rLIKEs in their speech), while older boys' usage of the discourse marker rose marginally, by only 4%.

Unlike Schweinberger (2011), D'Arcy (2017) found gender differences in the usage of clause-medial LIKE in Canadian English as well. In TEA, the corpus she used, men appeared to be the leaders, with the gender effect manifesting across all major syntactic contexts. D'Arcy shows that if the frequency of like in each syntactic position is plotted over apparent time, the gender distinction within each position is narrow or non-existent initially. Subsequently, as the frequency of like increases, so does the gender distinction, which then holds constant over apparent time. A similar trend is evident in Levey's (2006) data as well. Although he directly compared the usage of the discourse particle only in the DP and nP contexts, the effect is obvious: for both younger and older boys, the discourse particle before DP or nP accounts for 42% and 37% (respectively) of all the rLIKE tokens they produced. For the girls, the proportions are lower (25% for younger group and 20% for older group). Notably, the younger girls appeared to use more discourse particles than discourse markers, but the preference changed sharply by

the age of 10-11. Overall, Schweinberger (2011) claims that "in five of the eight regional varieties, it is clause-medial LIKE which is gendered – none of the other forms exhibits such a consistent pattern" (p. 393).

The situation with the quotative LIKE differs with regard to the variety of English as well. Tagliamonte & Hudson (1999) investigated the use of quotatives (including quotative LIKE) by British and Canadian youth and found different trends:

> "In Britain, where it is found more frequently overall (18%) and across many more speakers (i.e. more diffused), speaker sex is statistically significant to its appearance and females are clearly in the lead. In Canada, on the other hand *be like* is used less frequently (14%) and by far fewer of the speakers and speaker sex is not statistically significant, although females show a preference for it" (p. 161).

Finally, there is evidence that clause-final LIKE (in the varieties of English in which it is frequently used, such as Irish English) is also subject to gender-related conditioning. Contrary to expectations, this discourse element, which has traditionally been speculated to be characteristic of older males from rural areas (Nestor, 2013), may actually not be exclusive to this group. Schweinberger (2012) did not find significant gender differences in the use of clause-final LIKE, and later argued that it may be a result of the overt stigmatization of this form interfering "with general trends observed in the linguistic behavior of women" (Schweinberger, 2011, p. 393), meaning that the female lead (Labov, 2001a) is unlikely to be observed if LIKE is overtly stigmatized as a vernacular element. Corrigan (2015), however, found that women, who constituted 43% of the sample population, were actually in the lead, having produced 57% of all clause-final LIKEs, and further analysis showed that it was younger women who were responsible for this pattern and "innovating towards the traditional (male) vernacular variant" (p.

59). Based on these two studies, it can only be concluded that the stereotype of clause-final LIKE as a male speech element is not supported by evidence in contemporary Irish English.

Style and Pragmatics. Jucker & Smith (1998) proposed a distinction between reception and presentation discourse markers, describing LIKE as the latter and indicating its core function to be “flagging linguistic expression, for example, words, phrases, clauses or entire utterances, as less than literal representations of the thoughts they are intended to represent” (p. 191). Crucially, they demonstrated that presentation markers, of which LIKE was by far the most frequent, were more likely to be used in a conversation between friends rather than strangers; this pattern was true for LIKE specifically as well. Fuller (2003) investigated the usage of discourse marker LIKE in interviews and casual speech and found higher frequency of LIKE usage in interviews, the more formal context. In her interview data, LIKE was used "to focus on salient information, qualify contributions, and introduce examples" (p.370), and interviewers used them to step "into a more conversational pattern of discourse" (p. 371), especially if interviewees were reticent or, conversely, if a good rapport was established. These findings suggest that style and familiarity between interlocutors may be affecting overall frequencies of rLIKE, although it is not known whether these factors apply uniformly across different varieties of English (both studies were conducted with U.S.-based participants) and/or across different functions and syntactic positions of rLIKE.

Non-Native Speakers and rLIKE Usage. The usage of rLIKE by NNSs has been much less systematically researched, and only certain types of speakers have been targeted as large corporal data are not readily available (many of the existing learner corpora contain data from not sociolinguistic interview but specific tasks). There are two distinct types of NNS participants of rLIKE research: immigrants in English-speaking countries, usually with little formal training

in English (Diskin, 2017; Diskin & Regan, 2015; Nestor, 2013; Nestor & Regan, 2015), and NNSs in academic settings, usually graduate students and faculty members on college campuses (Fuller, 2003a; Liao, 2009; Liu, 2016). Interestingly, there is a geographic distinction as well: Most of the studies investigating rLIKE usage by immigrants were conducted outside of North America, in Ireland or in the UK, while most of the studies in academic settings were carried out in the U.S. Another difference is that researchers working in academic settings, unlike their colleagues working with immigrants, typically attempt to ensure sufficient exposure to spoken language and, therefore, to the vernacular element in question, prior to data collection. Two years of stay in the target language environment is the most common threshold for variationist research with NNSs in general (Drummond, 2011; Hellermann & Vergun, 2007; Schleef, 2013), and this is the often-chosen selection criterion used in rLIKE studies as well.

Usage of rLIKE by Immigrants. Nestor and Regan (2015) studied Polish immigrants in Ireland whose length of stay ranged from a minimum of one and a half years to over three years and discovered a clear trend: non-native speakers in the rural setting produced an overall high number of clause-final LIKE tokens and a lower number of clause-medial tokens, while their urban peers demonstrated the opposite trend. Age was also found to be a factor (independently of the setting): overall, younger speakers were more likely to prefer clause-medial LIKE, while adults tended to use clause-final LIKE frequently. Gender was not considered as a factor, although some evidence from the earlier study with a subsample of the same population (Nestor, 2013) allows for speculation (and speculation only, due to extremely small sample size) that non-native speakers may be sensitive to gender preferences: the clause-final LIKE may be associated not only with rural Ireland, but also with men, which could explain why only male speakers from the rural area in the sample demonstrated robust acquisition of clause-final LIKE (however, see

contradicting evidence for native usage in Corrigan (2015) and Schweinberger (2012)). Nestor, Chasaide, and Regan (2012) analyzed the same data set using both quantitative and qualitative techniques and found that, while as a group, NNSs exhibited patterns that "corresponded broadly" (p. 349) to native Irish patterns, there was significant interspeaker variation. Analyzing speaker data case-by-case, they speculated that this variation could have been caused by a number of factors, including the amount and type of input received (e.g., some of the study participants were exposed to American English prior to their arrival to Ireland), level of English proficiency (there was an apparent discrepancy between higher self-reported levels and the speech produced during the interviews by some participants), and the type of identity a given speaker was attempting to construct.

Diskin (2013, 2017) studied the usage of LIKE by Chinese and Polish immigrants in Dublin and found no correlation of frequencies with self-reported proficiency but a weak correlation between rLIKE frequencies and length of residence (LOR): While those immigrants whose LOR was only one to two years produced less than one token of rLIKE per 1,000 words, after three to four years of residence that number stabilized at the rate of nine to ten tokens, which was very similar to the rate exhibited by native speakers (NSs) (10-12 tokens per 1,000 words). However, the differences between groups were found in the syntactic placement of rLIKE and its pragmatic functions. Thus, NSs of Irish English used clause-final LIKE more frequently, while NNSs used more approximators. The hedging (mitigator) function appeared to be exclusively native, although it is unclear from the study whether NNSs did not hedge at all or used other discourse elements instead of LIKE for that purpose. While no overall gender effect emerged, females in all groups appeared to use LIKE as a filler (hesitation marker) more frequently than males. Working with the same data set, Diskin and Regan (2015) found type of

migrant (distinguished based on the primary goal of immigration) to be a factor affecting rLIKE usage. While non-native speakers in their sample were generally reluctant to use clause-final LIKE, the speakers most invested into their new Irish identity (the so-called "chain" and "economic" types of migrants) used it with higher frequencies than "cultural" migrants who possibly "did not aspire to 'sound Irish', but rather to sound unmarked, 'normal' and 'standard'" (p. 170). As summarized by Forsberg Lundell and Bartning (2015), "the avoidance of the 'Irish variable', clause-final like, by the NNS[s], shows a sensitivity to varietal variation and displays evidence of identity construction" (p. 10).

Truesdale and Meyerhoff (2015) investigated the use of "focuser LIKE" (the term that encompasses both discourse marker and particle, in D'Arcy's terms) and quotative LIKE among Scottish and Polish teenagers (14 to 17 years old) in Edinburgh, Scotland. The NNS group in the sample was highly diverse in terms of their LOR (varied between four months and four years) and time spent learning English (varied between nine months and seven years). Possibly due to this diversity, the overall frequencies of rLIKE usage between NSs and NNSs were found to be "very different overall" (p. 19), but the exact numbers, unfortunately, were not reported. However, the researcher compared the distribution across (mostly) pragmatic functions (lexical focus, word finding, clause-final, corrective, quotative) and found striking similarities between groups (e.g., 68% of native and 69% of non-native tokens were those of focuser LIKE, 15% and 16 %, respectively, of word search LIKEs, etc.). They also conducted analysis of how certain linguistic and social factors could affect rLIKE usage and found that for L1 Polish speakers, only speaker gender, a social factor, was a significant predictor of focuser LIKE usage (with females in the lead), while the nature of the preceding constituent, a linguistic factor, was selected as a single significant factor for the native speakers. As gender was not a significant factor affecting

native usage of focuser LIKE, the researchers interpreted the finding as evidence that "the teenage girls are moving towards the norms of their Edinburgh-born peers ahead of the boys" (p. 24).

Lee (1999) investigated the usage of three discourse markers (*LIKE*, *you know*, and *I mean*) by three generations of Korean immigrants in the United States and found that LIKE was favored by the immigrants with early age of arrival and those born in the U.S., while those who immigrated as older adults preferred the marker *you know*. This age-related pattern, however, appeared to be also gendered: Women, unlike men, did not show a steady decrease in discourse marker use in general or in marker preference, and, as a group, showed higher LIKE frequencies overall (21.54 tokens per 1,000 words vs. 17.20 for men)[5]. Generation-wise, the most striking differences were found between those born in the U.S. (Generation 2) and those who immigrated as children (Generation 1.5), on the one hand, and adult immigrants (Generation 1). Generation 2 speakers produced 20.47 tokens of LIKE per 1,000 words, for Generation 1.5 the number was even higher, 25.25, while Generation 1 speakers used LIKE only at a rate of 8.83. Unfortunately, no native speaker data from the place and time of Lee's (1999) data collection are available for comparison.

Hellermann and Vergun (2007) studied beginning adult learners of English at a low-cost community college adult English for Speakers of Other Languages (ESOL) course for immigrants and refugees for five years, attempting to determine the influence of LOR, level of proficiency, self-reported use of English outside of classroom on the discourse marker use (the focal markers were LIKE, *you know* and *well*). Their results showed that "more proficient students, those in the upper-level classes, use more of the focal discourse markers and they are

[5] Lee (1999) presents frequencies per 10,000 words, but, as he also reports total number of words, it was possible to recalculate the frequencies for better comparability with other studies.

also the students who it appears are more acculturated to English-speaking cultures of the U.S." (p. 176). However, these results were based on scarce data, because only 98 discourse marker tokens were extracted from 8,802 speaker turns. Among these, 63 tokens were those of LIKE, and their pragmatic functions were determined to be lexical focus (39), loose interpretation-approximation (20), and exemplifiers (four tokens).

To summarize, non-native speakers, while not always matching their native-speaking peers in overall frequencies of rLIKE usage, may be sensitive to such "socially salient" constraints, such as speaker gender. In some cases, however, such as in case of clause-final LIKE in Irish English, gender may be only a perceived constraint, as studies with native speakers (Corrigan, 2015; Schweinberger, 2012) fail to show gender differences. Furthermore, NNSs can be also sensitive to the native patterns of distribution across pragmatic functions and syntactic positions of rLIKE. There is also evidence that identity that can be indexed through rLIKE usage, both by overall frequencies and by its usage in certain syntactic positions. Age of arrival, length of stay in the target language country, and level of English proficiency may all be factors involved in the frequencies and patterns of rLIKE usage exhibited by NNSs in immigrant settings.

Usage of rLIKE in Academic Contexts. Out of six known studies of rLIKE usage by NNSs in academic contexts, two focused on L1 German speakers with significant exposure to American English in the target language environment. Müller (2005) examined discourse markers in the Giessen-Long Beach Chaplin Corpus, which included recordings of non-native English speakers (mostly L1 German speakers) during study abroad in Long Beach (California, USA). Davydova and Buchstaller (2015) studied the use of quotative markers by L1 German speakers of English in the Mannheim Corpus of German English (MaCGE), which contains

recordings of students "enrolled in degree programs [in Germany] with the highest uptake of study abroad programs" (p. 445). Both of these studies included rLIKE as one of several target elements (discourse markers or quotatives).

Davydova and Buchstaller (2015) investigated how different social and linguistic factors (mimesis, content of the quote, grammatical subject, tense/aspect, exposure to English, and speaker sex) affected the choice between multiple available quotatives (e.g., *say, think, be like*, etc.). Importantly, they took into account the overall non-nativelikeness of speech they were working with and included in the analysis not only the target-like form *be like*, but also such variants as [zero] *like* (interpreted as possible influence of German verbless structure), *say like, feel like, think like, know like*, and other (*sing like, tell like*, etc.) used to introduce a quote. The results indicated that only high-exposure learners (those who participated in study abroad programs) demonstrated acquisition of major probabilistic constraints on quotative LIKE usage, such as mimesis or content of the quote (thought vs. speech). It was also found that BE LIKE and its template forms were used by females at much higher rates than by males. Among the males, "*say* reigns supreme and *be like* trails in fourth place" (p. 458). The authors conclude that their findings "suggest that high-exposure German EFL learners aim–and manage successfully–to replicate the grammar that governs the native speakers' quotative usage" (p.465).

Müller (2005)'s study focused on four discourse markers (*so, well, you know*, and LIKE) and found considerable differences between the frequencies of three of them (*so, you know*, and LIKE) between NSs and NNSs. As for LIKE usage (both as a discourse marker and a quotative), certain trends emerged. Thus, both NSs and NNSs were more likely to use LIKE when talking to a friend than to a stranger, but this effect was more pronounced for NNSs. As a specific task was used for data collection, and those speakers who were narrating a movie to a partner used more

quotative LIKEs, a finding that applied to both NSs and NNSs, and NNSs in this role also showed higher frequencies of LIKE introducing examples or explanations. Among other factors influencing frequencies of LIKE in NNSs' speech were contact with native speakers (measured as time spent in and English-speaking country) and the influence of a particular variety of English: Those who spent more time in the U.S. than in Great Britain used LIKE more frequently across all pragmatic functions. It must be noted, however, that "all Americans used discourse marker *like* at least once, while this was the case for only 44 out of 77 Germans" (p. 230), even though the study included the usages of LIKE as a preposition, a conjunction, and as part of general extenders, such as "something like that" (usages not considered remarkable by (D'Arcy, 2017)). This means that 43% of the NNS participants did not use LIKE at all.

Given the importance that exposure to native speech appears to have with respect to the frequencies of LIKE usage, it could be expected that English learners in a foreign language setting who have never been in an English-speaking country would not use LIKE in its remarkable functions, e.g., as a discourse marker, at all. However, Algouzi (2015) found LIKE to be the most frequent discourse marker (out of three, the other two being *so* and *you know*) in the Saudi subcorpus of Louvain International Database of Spoken English Interlanguage (LINDSEI). The corpus included sociolinguistic interviews with third- and fourth-year university students aged 20–25 years from multiple universities and colleges in Saudi Arabia, and many of the students reported studying abroad prior to data collection. Out of 50 speakers, 41 used LIKE at least once, and the average frequency of LIKE as a discourse marker was 3.66 tokens per 1,000 words. Furthermore, "among the speakers in the S[audi]L[earner]C[orpus], there were some participants who lived abroad for years but used fewer instances of like than those who had never been abroad" (p. 217). The author even checked the English textbooks used in Saudi Arabia and

found that LIKE as a discourse marker or a quotative was not introduced in any of them. He concluded that “even though *like* is neither introduced in Saudi English language textbooks nor taught, Saudi English language speakers are probably able to acquire it through their exposure to the media and through their interaction in English with their peers” (p. 218).

The remaining three studies of rLIKE usage all involve graduate students and faculty members at U.S. college campuses. Fuller's (2003a) study of discourse marker use (focus markers included *well, oh, y'know*, LIKE, *I mean*) involved recording native- and non-native speaking graduate students and faculty members in three contexts: interview conducted by a NS, elicited narrative (narrating a picture book), and casual conversation (with close friends or family members). Narratives did not elicit many discourse marker tokens, so the author focused on comparing the interview and casual contexts. She found that NSs were more likely to use LIKE during interviews (8.3 vs. 7.3 tokens per 1,000 words), while the opposite trend was found in NNSs’s speech (3.4 tokens in interviews vs. 4.2 in conversations). However, interspeaker variation was high in both groups across both contexts: from 4.3 to 14.4 among NSs and from 0 to 5.8 among NNSs, which indicates that all NSs but not all NNSs produced LIKE tokens at all. Fuller noted that NNSs behaved native-like (in terms of using more or less frequently in the same context as NSs did) with regard to *oh* and *well* but not with *y'know,* LIKE, and *I mean*. According to her, "the non-native speakers continue to use D[iscourse]M[arker]s to modify and focus in conversations with people with whom they already share background knowledge, while the native speakers reserve such negotiations for interactions in which they need to create common ground" (p. 206).

Liao (2009) researched the ways in which six Chinese teaching assistants (TAs) used discourse markers (*yeah, oh, you know,* LIKE, *well, I mean, ok, right*, and *actually*) when

teaching (leading a discussion) and when being informally interviewed by a fellow NNS. The results showed that TAs used more discourse markers during informal interviews, which, as the author speculates, may be, in addition to the differences in formality of the context, a result of the fact that classroom speech observed was heavily planned beforehand (all of the TAs prepared detailed lesson plans, and one of them even wrote her notes in full sentences entirely). As could be expected, LIKE was among the markers least frequently used in the classroom. In general, Liao discusses the usage of discourse markers by her participants in terms of conscious identity construction: One of her focal participants, Iris, appeared to be "very concerned to construct a professional persona in English" and, while she "noticed native speakers use D[iscourse]M[arker]s such as *you know*, *like* and *well* a lot, she expressed no intention to follow those "kouyu" (colloquial words)" (p. 1325). Statistical tests Liao performed proved that Iris, indeed, used discourse markers much less frequently than other participants in the study.

Liu (2016) also studied the usage of various discourse markers, including LIKE, by ten Chinese and five American graduate students on a U.S. campus. The NNSs in the sample were divided into two groups: new arrivals and non-new arrivals (LOR over than 10 months). Based on the fact that LOR strongly correlated with exposure to native English speech, Liu called these groups "high-exposure" and "low-exposure". Members of low-exposure group did use some discourse markers, but none of them used LIKE. Those in high-exposure group, however, used it at a rate of 4.1 tokens per 1,000 words. The NSs' rate was 8.7 tokens. NNSs differed from the NSs not only in overall frequencies but also in terms of the pragmatic functions: While NSs employed a large range of pragmatic functions (searching for appropriate expression, marking approximate number or quantity, introducing an example, introducing an explanation, marking

lexical focus), NNSs used LIKE almost exclusively to either search for appropriate expression or to mark lexical focus.

Overall, studies conducted in academic contexts, just like studies with immigrants, demonstrate that length of residence and exposure to native speech are important but not categorical predictors of LIKE usage. NNSs with high exposure may abstain from using LIKE while learners with no exposure, as shown by Algouzi (2015), may be frequent users. Most researchers find pragmatic differences in NS and NNS usage of LIKE, but NNSs are capable of acquiring probabilistic constraints on rLIKE usage and may be sensitive to regional differences between varieties of English they are exposed to (Davydova & Buchstaller, 2015; Müller, 2005). Liao's (2009) study demonstrates how international TAs use discourse markers, including LIKE, as an identity construction tool, which supports the findings of immigrant studies (Diskin & Regan, 2015; Nestor et al., 2012).

This literature review shows that NNSs are sensitive enough to rLIKE to acquire it and, while not necessarily matching native speakers in frequency of rLIKE tokens per certain number of words, even approximate native patterns, that is, demonstrate the acquisition of probabilistic constraints on its distribution. However, the existing literature does not address the issue of social meaning(s) of rLIKE or stylistic constraints on its usage. Are NNSs aware of the fact that rLIKE is socially salient as a non-prestigious, in some ways stigmatized vernacular element, and if yes, why would the pick it up? A possible answer comes from mounting evidence in variationist research, not with regard to LIKE so far, that socially salient vernacular variants (e.g., *ne* deletion in French, *t*-glottaling in British English) may possess important, though largely not overtly and consciously recognized, symbolic value for NNSs. as it may be strongly associated with native-speakerness (“Frenchness” or “Britishness”) and thus boost its usage,

especially, although counter-intuitively, in more formal contexts where NNSs would prefer to present the best version of their L2-speaking self. Thus, in her study of glottal replacement among Punjabi immigrants in London, Sharma (forthcoming) describes how first-generation immigrants use this vernacular feature, which is associated with informality and lower social status, during interviews. She argues that they do so "to signal a polite, formal accommodation to a British style, possibly to signal legitimacy or competence to a community outsider affiliated with a British institution", which she interprets as the orientation of variation "to an 'us-them' contrast that is more salient to them than internal class hierarchies in British society, indexicalities that many of them have very limited social access to" (p. 10). Sharma also cites a number of studies which found similar trends of "stylistic reinterpretation" (Labov, 2001b) with regard to (aN)-raising in Tehran (Modaressi-Tehrani, 1978), ə-lowering in Bergen (Kerswill, 1994), glottal replacement and other vernacular forms in London (Cheshire, Kerswill, Fox, & Torgersen, 2011; Sharma, 2016). Regan, Howard, and Lemée (2009), discussing the changes in the French interlanguage of L1 English students upon their return from study abroad, note:

> "Interestingly, although their experience in France confirms their hypothesis that *ne* is retained more in formal style, after the year abroad, style makes slightly less of a difference to deletion rates than before in the speech of these learners. The L2 speakers delete more in monitored style in Time 2 than in Time 1" (p. 72).

It is interesting, therefore, to investigate whether a similar process of overusing rLIKE in order to project an American identity could be observed among my non-native-speaking participants.

Remarkable LIKE as a Carrier of Social Meaning

Remarkable LIKE is omnipresent in native English speech, especially among adolescents and young adults. It is also socially salient, unlike some other ubiquitous vernacular elements. Experimental studies show that native speakers of English notice it in speech (Maddeaux & Dinkin, 2017) and may judge the same speakers differently on metrics of friendliness and intelligence depending on whether they use rLIKE in their speech (Buchstaller, 2006; Dailey-O'Cain, 2000; Hesson & Shellgren, 2015). Additionally, many stereotypes about rLIKE exist and are widely circulated in the media (D'Arcy, 2007, 2017).

While there is ample evidence that non-native speakers of English may use rLIKE (Chloe Diskin & Regan, 2015; Fuller, 2003a; Liao, 2009; Müller, 2005; Nestor & Regan, 2015), very little is known about whether rLIKE is equally salient to them and whether non-native beliefs about and attitudes towards rLIKE differ from those that native speakers possess. The main purpose of this part of the chapter is to explore these questions. I turn first to a general discussion of the perception of sociolinguistic variation, then clarify the terms I chose to use in this dissertation to discuss perception (beliefs and attitudes), and finally summarize existing research on native beliefs about and attitudes towards rLIKE.

Social Meaning and Sociolinguistic Perception

Sociolinguistic variables offer speakers a choice between two or more variants, one of which can be zero, in accordance with the principle of accountability proposed by Labov (1972). Different variants, in addition to being constrained linguistically in certain ways, may carry important social meaning, such as gender, membership in a socioeconomic class, ethnic and/or geographic origin, or more locally specific meanings etc., thus allowing speakers to index (consciously or unconsciously) certain types of identity. Overt or covert prestige can be attached

to a particular variant (Trudgill, 1972): For example, while women in Trudgill's study of Norwich speakers were more likely to use the standard velar nasal [ɪŋ] variant of variable (ing), associated with higher socioeconomic status and thus overtly prestigious, men were relatively more likely to use the non-standard form [ɪn], associated with masculinity and carrying covert prestige not immediately evident to an outside observer.

Sociolinguistic variants bear meanings not only for the speaker but also for the listener. This is straightforwardly evident for sociolinguistic stereotypes that are available for metalinguistic commentary by speech community members such as *ain't, y'all* (Niedzielski & Preston, 2000). For sociolinguistic variants that do not reach this salience threshold, experimental studies have been employed to explore their associated social meanings. Multiple studies have shown that listener perceptions, understood as "processes engaged when people are exposed to external stimuli, in this case linguistic material, and extract information from it" (Campbell-Kibler, 2010, p. 378) may differ depending on what social information about the speakers is available (see Campbell-Kibler (2010) for a review). For example, in a key study directly targeting the influence of social information on listener perceptions, Niedzielski (1999) asked listeners from the Detroit area to listen to some target words in recorded sentences and then select vowels that best matched those they heard in the sentences from a set of computer-resynthesized vowels. The speaker (a Detroiter) was presented as a fellow Detroiter to some listeners but as a Canadian from Windsor (a city located directly across the border from Detroit) to others. For words containing the diphthong /aɪ/ (as in *life, right*), the Detroiter realized the diphthong with a raised nucleus [əɪ]. This pronunciation is used on both sides of the Detroit-Windsor national border, but in public metalinguistic discourse it is a linguistic stereotype of Canadian English. Accordingly, participants who believed they were listening to a "Canadian"

exhibited greater accuracy in choosing raised-diphthong tokens as best matching the speech, than participants who had been told the speaker was from Detroit.

Later studies have shown that social information is such a powerful filter on listener perception, that it can be effective even when it is presented implicitly. Thus, Babel and Russell (2015) found that listeners tended to rate speakers as less intelligible if they were visually primed by photos of Chinese Canadians, while no such effect was observed when pictures of white Canadians were used as primes. Hay and Drager (2010) demonstrated that stuffed toys symbolically associated with certain countries (kiwi birds with New Zealand, kangaroos with Australia) influenced listener perception of vowels simply by being visually present in the lab for a few moments when a participant entered the experiment room, and then being taken away.

As we see from these examples, studies of sociolinguistic perception have largely relied upon (or have for convenience been constrained to) the one-to-one mapping of sociolinguistic variants to fixed and context-independent social meanings such as “Canadian” for raised /aɪ/ or “high status” for [ɪŋ]. This unidimensional view of social meaning has however been reconsidered by many researchers over the last two decades, as a “Third Wave” approach to sociolinguistics has developed (Eckert, 2005).

The Third Wave of Sociolinguistic Research and the Notion of Indexicality. In her seminal paper “Variation, convention, and social meaning”, presented at the Annual Meeting of the Linguistic Society of America, Eckert (2005) narrated the history of language variation and change research by dividing it into three overlapping ‘waves’. The difference between the waves lies in theoretical and methodological approaches to the study of variation. The First Wave originated with Labov's (1966) study of social stratification of English in New York, which established “broad correlations between linguistic variables and the primary social categories of

socioeconomic class, sex class, and age" (Eckert, 2005, p. 1). The Second Wave was focused primarily on investigating "more locally-defined populations" (p. 1) such as Eckert's own famous study of Jocks and Burnouts at a Detroit area high school (Eckert 2000), and studies in this vein widely employed ethnographic methodology. Within the Third Wave, "there has been an emerging focus on variation not as a reflection of social place, but as a resource for the construction of social meaning" (p. 1).

Describing this new approach, Eckert (2005) focused on the concept of a social persona, an iconic example of a social group or type:

> "When we [i.e., the public] think about the relation between variation and social groups, we don't generally identify individual variables. We have constructs in mind like Valley Girls, New York Jews, Mafiosi, Rappers, Southern Belles – persona types that constitute an ideological social landscape. The variables that characterize the varieties associated with these types do not themselves generally mean "Valley Girl, New York Jew" etc., but combine to produce those meanings. In other words, the meaning of variation lies in its role in the construction of styles, and studying the role of variation in stylistic practice involves not simply placing variables in styles, but in understanding this placement as an integral part of the construction of social meaning" (p. 24).

As Hall-Lew and Stephens (2012) pointed out, "the comment that someone "talks country" is readily interpretable among speakers of U.S. English, and Country Talk is frequently invoked in studies of folk linguistics (p. 257). The instant recognition potential of various social personae types, which is based on the indexical connections between stylistic practices manifested in speech and social meaning, allowed Campbell-Kibler (2009) to define social meaning as "social content tied in the minds of a given speaker/hearer to a particular piece of

linguistic behavior" and underscore that linguistic behaviors and social structures correlate not on their own, but because "speakers/hearers mentally connect them, whether consciously or unconsciously" (p. 136). Campbell-Kibler (2011) underscores that the linguistic stylistic practices Eckert is referring to are "fundamentally similar to nonlinguistic social practices such as the wearing of particular clothes" (p. 425) in that they index not only social categories, such as age or sex, but stances and other situation-based meanings as well. Therefore, "through constantly repeated acts of indexing, agents build their social selves and the societal structures they inhabit. At the same time, this iteration of use builds and maintains the indexical connections themselves" (p. 425). To refer to this "constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable" (Eckert, 2008, p. 454), Eckert proposed the term "indexical field".

In her study of speaker attitudes towards variable (ING), Campbell-Kibler (2009) tested the hypothesis that (ing) is interpreted by listeners not just on a unidimensional scale of formality-informality, but via a complex indexical field. To do this, she made an important methodological choice as she decided not to give to listeners any information about the speakers, explicitly or implicitly, so that they were able to profile speakers freely, based on their speech only, after which Campbell-Kibler recruited other (but demographically similar) listeners to judge the intelligence, education level and other attributes of the speakers. The free profiling part of the study conducted with focus groups allowed Campbell-Kibler to investigate the effects of not only of macro-social categories such as region, education and class, and traditional subjective attributes like intelligence, but also focus-group generated attributes that included social personae. In public discourse, these may have enregistered labels such as "soccer mom" or "frat guy"; the personae mentioned by Campbell-Kibler's participants included Valley Girls,

rednecks, stoners, metrosexuals, etc. In an earlier paper on how the [ɪn] variant of the (ing) variable may belong to two different socially constructed accents, "Southern" and "gay", Campbell-Kibler (2007) already demonstrated that both macro-social and traditional subjective attributes also interact with social personae. In conclusion, she underscored that "while these interconnections can be frustrating for researchers as we attempt to tease apart the strongest, most relevant connections for a given accent or linguistic variable, they are a fundamental aspect of sociolinguistic variation" (p. 55).

D'Onofrio (2015), continuing to focus on the complexity of the relationship between social information and sociolinguistic variables, investigated the impact of differently organized social information on listener perceptions. In addition to macro-social categories, such as region of speaker's origin, she used the concept of social persona and presented speakers as "nerds" or "Valley girls". Using a word-identification task with the TRAP vowel /æ/, the backing of which is an element of a chain shift currently in progress in California, she manipulated[6] the auditory stimuli for a word choice task in which social information about the speaker was either not given or presented in one of two forms: as a macro-social category ("from California" or "from Michigan") or as a particular social persona ("Valley Girl" or "Nerd"). Both the choices listeners made and their pre-choice eye movements were analyzed (the icons showing either shapes of the states or symbols for the personae were on the screen). Summarizing the findings, D'Onofrio underscored that listeners used persona-based information to shape their linguistic expectations at both pre-choice (automatic processing) and word choice levels, and that listeners "who thought they were hearing a specific *type* of Californian were led to categorize an ambiguous

[6] "Nine-step continua from each of the recorded TRAP tokens to respective LOT tokens (e.g. SACK to SOCK) were … created" (D'Onofrio, 2015, p. 245).

word as TRAP-backing to a significant extent, while their counterparts who thought they were hearing a Californian did so to a marginal degree" (p. 251).

Collectively, Campbell-Kibler (2007, 2009) and D'Onofrio's (2015) experiments underscore Eckert's (2005, 2008) claim that the relationship between social meaning and sociolinguistic variables is complex, context dependent, fluid, dynamic and multidimensional. Understanding the complexity involved in the way listeners extract social meaning from sociolinguistic variables is important for both designing a variationist study and interpreting its results, especially when it comes to variables that are socially salient, such as rLIKE. Because a listener may base his or her evaluation of a given form, or discourse element, on the type of social persona this element is associated with, a detailed analysis of such associations is warranted.

Language Ideologies, Beliefs, Attitudes, and Stereotypes: A Note on Terminology

Language ideologies can be defined as "sets of beliefs about language articulated by users as a rationalization or justification of perceived language structure and use" (Silverstein, 1979, p. 193; cited in Kroskrity, 2004).The scope of language ideologies includes "the full range of scholars' notions of ideology: from seemingly neutral cultural conceptions of language to strategies for maintaining social power, from unconscious ideology read from speech practices by analysts to the most conscious native-speaker explanations of appropriate language behavior" (Woolard & Schieffelin, 1994, p. 58). Kroskrity (2004) conceptualized language ideologies as grounded in social experience and multiple, "because of the plurality of meaningful social divisions (class, gender, clan, elites, generations, and so on) within sociocultural groups that have the potential to produce divergent perspectives expressed as indices of group membership" (p. 50).

The idea of language ideologies has direct implications for sociolinguistic research, especially in the area of discourse pragmatics. Native speakers acquire sociolinguistic variation simultaneously with learning the cultural norms and (usually sophisticated) social hierarchies of a certain community, and typically at a comfortable pace. For example, nobody expects a young child to be as well-versed in the community structure as adult members of that community. Importantly, native speakers learn "from scratch", as they are not born with any particular language ideologies. For non-native speakers, the process of acquiring sociolinguistic competence is much more complicated. There is the pressure of the expectations imposed by the new community to follow a myriad of unwritten rules, from using age- and style-appropriate greetings to understanding a local dialect that may sound very different from the variety of the language a newcomer is familiar with. Additionally, non-native speakers face multiple challenges adapting to the new environment in the linguistic and non-linguistic senses, while not necessarily considering this new environment as a "final destination", a permanent place to live in. In the model of learner investment into a language proposed by Darvin & Norton (2015), ideology is considered a crucial component, along with identity and (symbolic) capital that is the major source of affordances for learning and benefits that learners may or may not perceive as such. For example, a decision to use or avoid a particular word or a discourse feature, or even the decision to talk or to stay silent altogether may depend on whether the learner feels comfortable enough to take risks and potentially lose face, which, in turn, may depend on the difference or similarity in status between the learner and his or her interlocutor in a given context. The author describes this process as the learner constantly being positioned and repositioned by the systemic patterns of control operating at all levels.

In this study, I understand ideologies as a macro-level category that includes everything speakers "know" and "think" about human language in general, particular language or languages they speak, any linguistic units or discourse elements, linguistic behavior of native and non-native speakers of different age and gender and representing different social personae, etc., regardless of whether they are aware of possessing this "knowledge" and having these "thoughts", which aligns with Kroskrity's (2004) understanding. While this is not the only possible way of understanding language ideologies (e.g., De Costa (2011) uses the term "beliefs" to refer to this overarching construct of everything above and below the level of speaker's awareness), this is a common in attitudinal research and convenient way to interpret the concept of language ideologies and ensure a clear terminological distinction between the terms "ideologies" and "beliefs".

"Beliefs" and "attitudes", however, are the two terms I use most frequently and distinguish for methodological purposes. Beliefs include all "knowledge" and "thoughts" that speakers are aware of and are able to formulate in some way. For example, if a participant says that using rLIKE makes people sound stupid, I understand it as his or her belief. Crucially, individual belief systems are situated and context-dependent, and thus not necessarily structured and logical. In other words, a person may present different or even contradicting beliefs about the same phenomenon. In addition to that, while no two individual belief systems are identical, there is a large portion of shared beliefs in the individual systems of beliefs among the speakers belonging to the same speech communities, including the large community of people speaking a particular language. Stereotypes are the "public domain property" in a belief system; speakers are aware of stereotypes as of something that "most people think", even if they do not agree with it. Finally, attitudes constitute the below-the-level-of-awareness part of ideologies an individual

possesses. Attitudes, unlike beliefs, cannot be investigated using interviews and questionnaires and require special techniques, such as matched-guise methodology. While I have not seen these exact definitions in the literature, it seems that this distinction is implied in many sociolinguistic works. For example, in the opening sentence of his chapter on attitudes in "The Handbook of Language Variation and Change", Preston (2013) writes: "The study of language attitudes focuses on the linguistic clues that both guide a hearer to a speaker's group membership and trigger the hearer's beliefs about the group" (p. 157). Thus, the author implies that beliefs are something at the level of awareness, while attitudes work below that level.

Even though I make this terminological distinction in this dissertation, I still call the techniques I used to investigate beliefs "attitudinal interview" and "attitudinal questionnaire" and I refer to the large body of relevant research as to "attitudinal research", simply because these names are traditionally used in the field.

Popular Beliefs about rLIKE

Drawing upon multiple attitudinal studies, remarks made by both authors and participants in various research papers, and mass media discourse, D'Arcy (2017) summarized the myths and "an intricate and multifaceted lore surrounding LIKE" (p. 176). She presented them as a six-component belief complex. The list presented by D'Arcy is the following:

- "*Like* is just *like*, that is, there is one *like* that is recycled repeatedly";
- "*Like* can be used anywhere in a sentence";
- "*Like* began with the Valley Girls";
- "Only young people, and adolescents in particular, use *like*";
- "Women say *like* more than men do";
- "*Like* is meaningless; it simply signals a lack of articulacy".

Some of these beliefs are socially neutral in that they only reflect the lack of linguistic expertise on the part of ordinary speakers and the nonsaliency of syntactic constraints. Other beliefs, however, portray LIKE as a popular culture phenomenon, as a signature feature of youth's speech, as a feature of female speech, and as a symbol of speaker's inarticulateness. Some of these beliefs have certain truth value (e.g., even though not only young speakers use rLIKE, they use it with much higher frequency), while others may be inaccurate (e.g., as D'Arcy showed, certain functions of rLIKE are, in fact, more frequently used by male rather than female speakers).

While D'Arcy did not directly address beliefs about racial differences in the use of rLIKE, and no such beliefs are widely discussed in mass media, the association between extensive use of rLIKE and whiteness may be inferred from the type of social persona frequently blamed for dissemination of rLIKE – the Valley Girls (the names more recent and overlapping in the social meaning behind it include "basic white girl", "Kardashian", "Becky"). The Valley Girl "is a popularly recognized female persona that is typically white, feminine, affluent, materialistic and superficial" (D'Onofrio, 2015, p. 243). Other types of white people can be associated with LIKE as well: For example, Bucholtz (2011) described the usage of quotative LIKE as a signal of "preppy whiteness" in a high school in California.

As is evident from the sheer number of myths surrounding LIKE, it is trendy, both in academic research and popular media concerned with the inevitable demise of the English language, to claim that LIKE is the invention of careless youth who do not speak properly. D'Arcy (2017) found that one of the instructions on the website wikiHow.com, whose mission is "teaching anyone in the world how to do anything", was entitled "How to Stop Saying the Word *Like*" (p. 29). In my own web searches, I found that another instruction, found on

OnlineCollege.org, gave advice on how to "stop saying *like* and immediately sound smarter". I also found a letter a concerned mother wrote to *The Guardian* newspaper to complain that "her daughter sounds stupid and uneducated because she uses the word 'like' all the time", and the advice columnist, Mariella Frostrup, responds that "it's just the world she lives in," the world in which "from Peking to Patagonia, culture (unless it's popular) and vocabulary (unless it's abbreviated) have been swept from our lives in the international language of mediocrity" (Frostrup, 2014). As summarized by D'Arcy (2017), "despite overwhelming empirical evidence of widespread use, not only historically but also regionally and socially, LIKE is not liked" (p. 42). The stigma attached to rLIKE may even affect people's hireability, as demonstrated in a study based on mock job interviews (Russell, Perkins, & Grinnell, 2008). As the study found, "both professionals and students [who played the role of interviewer in the study] were least likely to hire or recommend interviewees that used the word "like" compared to "uh" or the control conditions [without filled pauses or LIKEs]. Interviewees using "like" were also perceived as unprofessional by both students and professionals" (p. 116).

Fox Tree (2007) investigated folk notions of several discourse-pragmatic features, including rLIKE. Over one hundred undergraduate students, all native speakers of American English, responded to her questionnaire that targeted self-assessment of use, history of discussing use, attitudes (measured by asking the participants to choose a statement they agree with most), and folk notions of rLIKE meaning (investigated by using open-ended questions). Most of the participants claimed that they use rLIKE more with friends than with authority figures, and 77% of respondents reported trying to avoid rLIKE and pause fillers (*um* and *uh*) in their speech. Furthermore, "*like* was described as meaning nothing or meaning the same as some other marker by 51% of respondents", and "35% described like as being used out of habit" (p. 305).

Summarizing the results of her study, Fox Tree noted that while the participants were able to formulate what other markers studied (*um, uh, you know*) meant, "*like* defied definition" (p. 306). To dig deeper into the way laypeople understood the meaning of rLIKE and other discourse markers, she conducted an in-class experiment in which she gave transcriptions of naturally produced speech excerpts either in the original form or with substituted discourse markers (e.g., *you know* inserted into an utterance instead of *like*) and asked them to choose from a set of possible interpretations. After analyzing student responses, Fox Tree concluded that "even if laypeople cannot articulate precisely what like means, they do have a sense for how it can be used. They recognize that discourse markers cannot substitute for each other without changing meaning" (p. 307).

Attitudes towards rLIKE

The openly shared beliefs about rLIKE, as described earlier, are overwhelmingly negative. rLIKE may be discussed with regard to "the overarching and timeless gestalt that the language is deteriorating" (D'Arcy, 2017, p. 175), it can be attributed to inarticulate youngsters, and young women in particular. However, experimental methodologies that rely on listeners' perceptions that are fluid and partially unconscious rather than on overt questions in surveys and questionnaires allow for painting a more nuanced picture of the attitudes towards rLIKE. While no existing studies investigated non-native attitudes towards rLIKE specifically, native attitudes have been quite extensively researched.

Dailey-O'Cain (2000) used both a questionnaire and a matched-guise experiment to investigate the attitudes towards a "focuser LIKE" and a quotative like among forty native speakers of English, aged either 18-30 or 45-50.

> "Two example sentences containing *like* were provided, one containing focuser *like* and one containing quotative *like*, and informants were asked to fill out a short anonymous questionnaire in which they were asked to comment on whether they associate *like* usage with younger people or older people, with men or women, what they think of *like* in general, whether they perceive a distinction between focuser *like* and quotative *like*, and whether or not they use it themselves" (p. 68).

The results indicated that the majority of both younger and older informants believed that LIKE is more frequently used by young people, and both male and female informants reported that women use LIKE more often (although six informants believed in the equal distribution). Twenty-nine of the informants stated that they disliked rLIKE (quotative and focuser equally), mostly because "it makes people sound uneducated and lazy" (p. 70). Younger people and females of all ages reported that they themselves used rLIKE "often" or at least "sometimes".

For the matched-guise experiment, Dailey-O'Cain used "eight one-minute pieces of naturally-occurring speech" (p. 71) from four speakers (one young (17-19 y.o.) speaker of each gender, one middle-aged (33-34 y.o.) speaker of each gender). Each speaker contributed two monologues, one containing 12-15 uses of rLIKE and one from which all instances of LIKE were digitally removed. Informants were instructed to guess the age of and evaluate each of the eight presumably different speakers "on a scale from one to five for nine bipolar traits: 'attractiveness', 'cheerfulness', 'educatedness', 'friendliness', 'interestingness', 'intelligence', 'reliability', 'responsibleness', and 'successfulness'" (p. 72).

Predictably, *like*-guises were unanimously perceived as produced by younger speakers. More interestingly, while the general stereotypes of rLIKE-users as less educated and less intelligent were confirmed (though only the first at the statistically significant level), the *like*-

guises were also perceived as produced by more attractive, more friendly, and more successful speakers, and all three findings were statistically significant. A particularly fascinating finding was that "when the younger speakers used *like*, they were perceived as more interesting than when they did not, and when the older speakers used *like*, they were perceived as less interesting than when they did not" (p. 73). The results allowed Dailey-O'Cain to conclude that the use of rLIKE mostly appeared to positively affect the perceptions of the speaker's solidarity-oriented traits (such as friendliness, attractiveness, cheerfulness), but negatively affect the perception the same speaker's status-oriented traits, such as level of education or intelligence.

Buchstaller (2006) conducted a matched-guise experiment and administered a survey to reveal the attitudes of British English speakers from England, Wales, Scotland and Ireland (in age groups similar to those in Dailey-O'Cain's (2000)study) towards quotative LIKE (henceforth, qLIKE). While 93% of the respondents associated qLIKE with younger speakers, no clear consensus was achieved with regard to gender and class: almost 60% used the "I don't know" response option. A slightly different list of traits was used to accompany the matched-guise task: "*calm – giddy; trendy/cool – old-fashioned; educated – uneducated; annoying – pleasant; British – non-British; animated – boring; intelligent – stupid; confident – non-confident; extroverted – introverted; professional – unambitious; glamorous – dull; popular – unpopular*" (p. 371). As a result, the use of qLIKE was significantly associated with speakers perceived as more giddy, animated, and trendy/cool, but less educated, ambitious, and pleasant. Additionally, the British informants perceived qLIKE as an American rather than a British feature. It must be noted, however, that the guises were presented to the informants in written form. Due to the lack of research on the differences in the perception of spoken vs. written sociolinguistic guises, it is unclear how much of the differences between Dailey-O'Cain's (2000)

and Buchstaller's (2006) results can be attributed to the manner of presentation of the stimuli rather than the regional differences.

Hesson & Shellgren (2015) used a matched-guise experiment to continuously measure "real-time listener evaluations of speech samples differing only by a single use of D[iscourse]M[arker]L[ike] using a dynamic motion-capture interface" (p. 154). The participants, sixteen undergraduate students, used a drawing tablet with two axes representing the scales of intelligence and friendliness (the center of the tablet screen represented a neutral judgment). While listening to each of the ten audio excerpts (each presented in either authentic version, or digitally manipulated to remove all instances of rLIKE), the participants moved the dot to represent their current judgment of the speaker along the scales. The presence of rLIKE, as expected, prompted the participants to judge the speakers as less intelligent, but also, contrary to predictions based on previous studies, as less friendly. The negative effect on friendliness, however, decreased over time. As the authors summarized, "listeners seemed to produce knee-jerk "*like* is bad" reactions across both social traits, but after processing several additional seconds of sociolinguistic data on the speaker, [they] only maintained negative reactions with respect to intelligence" (pp. 172-173).

The results obtained by Dailey-O'Cain (2000) and Hesson and Shellgren (2015) may serve as a valid reference point for research with native speakers residing in Michigan: Dailey-O'Cain's participants were native Michiganders from the Upper Peninsula, and Hesson and Shellgren recruited undergraduate students at the Michigan State University campus in Lower Michigan. However, the evidence for non-native speaker attitudes is limited to anecdotes. For example, explaining the unexpectedly low frequency of rLIKE usage by a highly proficient English speaker with high exposure to naturally occurring conversation, Liu (2016) noted: "After

the data analysis, I interviewed Xia about the use of the D[iscourse]M[arker] *like*. Xia said that she avoided using *like* on purpose whenever she spoke English, because she thought that the marker was stigmatized and she preferred a more formal English style" (p. 87). Liao (2009), discussing the usage and non-usage of rLIKE by her participants, Chinese teaching assistants (TAs) on an American college campus, speculated: "As TAs in the classroom, the participants were aware of their roles as authorities, and thus they might tend to avoid using certain D[iscourse]M[arkers]s in their speech due to the stylistic connotations of these DMs as very informal and colloquial (e.g. *like* and *you know*)." Unfortunately, the authors (Liao, 2009; Liu, 2016) did not pursue the question of how these NNSs came to know that rLIKE was stigmatized and generally avoided in formal situations, that is, whether they were explicitly told that by a teacher, mentor, or someone else they knew, or somehow came to that knowledge by observing native speaker behavior.

Conclusion

In this literature review, I summarized the existing research on remarkable LIKE, paying special attention to the studies conducted with non-native speaking participants. It is evident from this review that such research has been largely constrained to rLIKE usage, comparing the frequencies in native- and non-native speech and between lower- and higher-proficiency NNSs, investigating the patterns of using different pragmatic functions of rLIKE (i.e., as a filler, hedge, etc.), and determining social factors affecting the frequencies and usage patterns. The insights gained from this literature allow me to expect that international students who spent considerable time on a college campus, given their high proficiency level (ensured by the college selection process that involves passing standardized test score threshold) and exposure to native speech, will use rLIKE in their own speech. Length of residence is most likely to be an important

predictor of rLIKE frequency; it is difficult to make predictions with regard to gender due to conflicting findings of studies conducted with different participant populations. While there is limited evidence that that style and relationships between the interlocutor may play affect rLIKE usage, it does not allow for making clear predictions with regard to it for two reasons. First, if the study design does not enforce different levels of formality across tasks, it is impossible to reliably establish how a given participant interpreted the level of formality of the only interview with the researcher. More importantly, there is evidence that NNSs may demonstrate "inverse behavior" and supply more rather than less salient vernacular features if they interpret the context as formal and want to signal their Americanness.

While the body of research devoted to rLIKE usage in non-native speech is sufficiently large, such avenues of research as syntactic placement of rLIKE in non-native speech, NNSs' sensitivity to rLIKE placement in native speech, as well as NNSs' attitudes towards and beliefs about rLIKE, remain unexplored, thus necessarily establishing this study as exploratory. Attitudinal studies conducted with native speakers reveal that rLIKE may be have a different effect on listener judgments with regard to solidarity-based speaker characteristics (e.g., friendliness, attractiveness) and status-based characteristics (e.g., intelligence, educatedness), and the persona-based approach to interpreting the way linguistic variables index social meaning, central to the Third Wave of sociolinguistic research, appears to be a promising data analysis tool.

CHAPTER THREE

METHODOLOGICAL OVERVIEW

The primary goal of this dissertation is to explore the ways in which non-native-speaking international students interact (from both the production and reception points of view) with remarkable LIKE, an element of the local vernacular which was, presumably, unfamiliar to them prior to their arrival to the English-speaking environment. In order to fulfill this purpose, a wide range of research methods and techniques was employed, and a large number of non-native and native-speaking participants was recruited. Each of the tasks or experiments presented to the participants was intended to collect a certain type of data (e.g., usage of LIKE or attitudes towards LIKE), so those tasks (both their design and treatment of the collected data, including coding and analysis) will be described in the appropriate chapters. For example, I conducted two different interviews with each non-native speaker, one aimed at eliciting LIKE in unprepared speech and another prompting the participants to overtly discuss what they think about the word "like". The first of these interviews will be described in detail in the Methodology subsection of Chapter Four and the second in the Methodology subsection of Chapter Five; in both cases the methodologies will be presented after the review of the existing literature on the subject (e.g., usage of LIKE or attitudes towards LIKE).

In this part of the dissertation, I will summarize information relevant to all parts of the study: the description of context in which the study was conducted and overview of the procedure, so that the reader will know how and in what order different tasks were presented to the participants and how the procedures for native and non-native speakers differed, and, finally, the process of participant selection and participant profiles, as well as information about the tasks each participant took part in.

Context

Native and non-native speakers of English were recruited for this study from the Michigan State University campus. The state of Michigan is one of the top ten destination states for international students. In 2017, Michigan State University (MSU) hosted 6,858 international undergraduate students, which constitutes 13.7% of the MSU student body. Crucially, while welcoming a large number of international students, most of whom speak English as a second language, MSU is relatively homogenous with regard to the native variety of English these students will be exposed to: in the same year of 2017, 72.7% of the domestic undergraduate students at MSU were Michigan residents (Michigan State University, 2018).

Information about the languages international students speak is not routinely collected by the university, but the country of origin may serve as a proxy for that type of information. In 2016, China was the leading country of origin among all the international students at MSU, with 3,687 students enrolled as undergraduates. The Republic of Korea supplied 232 undergraduates, and 124 students came from Saudi Arabia. The remaining countries in the top-ten list were India (108), Taiwan (88), Malaysia (77), Thailand (36), Canada (35), Hong Kong (30), Indonesia and Angola (27 each).

Undergraduate students usually have a busy schedule that requires them to attend classes every day, which ensures international students' consistent exposure to native speech on campus even if they do not participate in any extracurricular activities or social events and prefer to socialize among the speakers of their first language. The main sources of native input for a socially inactive student would thus be their classmates as well as instructors and professors. Most of the native-speaking classmates, as demographics suggests, are Michiganders, so a steady supply of vernacular elements (including LIKE), used in a consistent manner, can be assumed. It

can also be assumed that international students have an opportunity to observe the differences in vernacular usage in more and less formal contexts. Most faculty members and teaching assistant treat the classroom as a formal environment and thus deliberately try to avoid vernacular elements[7], with, perhaps, the exception of the so-called academic discourse markers, or "lecturer's OK" (Levin & Gray, 1983; Schleef, 2008), which do not include LIKE.

Overview of the Study and Procedure

The study involved a combination of experimental and non-experimental tasks. Some of the tasks slightly varied between the groups of participants or were given to only one group. The principal difference between the way NSs and NNSs participated in the study was the medium of participation: While NSs completed all tasks online, using *Qualtrics* survey software, NNSs met with me (the researcher) in person. In this section, I will only outline the procedure and give a brief description of each task. Detailed information about the tasks may be found in the next section, Tasks and Materials; the flowchart version of the procedure is presented in Figure 3.1.

Regardless of the medium (online or offline), all participants were first given a consent form. The study was vaguely yet not deceptively described as "a research study of conversational English" and the participants were informed about their rights, including the right to revoke their consent at any point, and compensation (course credit for NSs and monetary compensation for NNSs). As I met with NNSs in person, I walked them through the form in addition to giving them time to read it and encouraged them to ask questions. NSs were asked to confirm their consent by pressing a button; NNSs were asked to explicitly verbally state their desire to proceed with the study. After that, the participants chose a pseudonym and filled out the *background questionnaire*, presented on paper to NNSs and online to NSs. The primary goal of the

[7] It applies to international teaching assistants as well, as shown in Liao (2009).

background survey was to collect demographic information about the participants (age, gender, place of birth) and their linguistic profiles (languages they and their families speak). The NNS version of the survey can be seen in Appendix A, the NS version of the survey administered via *Qualtrics*, can be found in Appendix B.

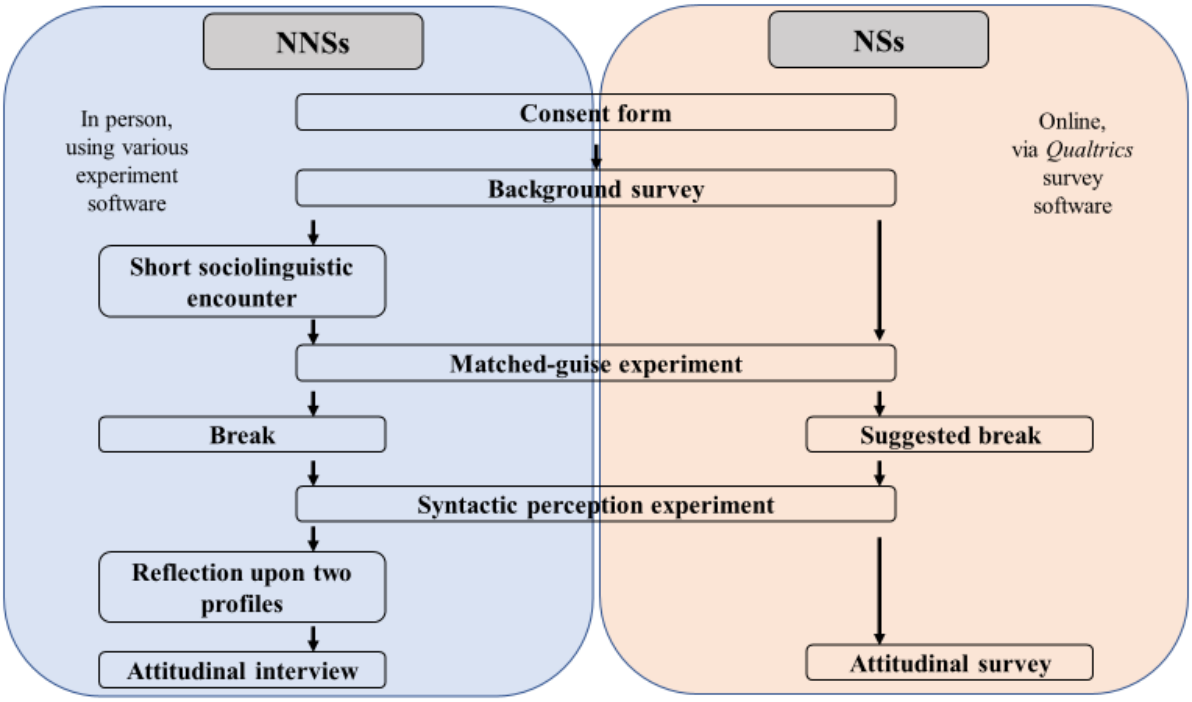

Figure 3.1
Data Collection Procedures for Native and Non-Native Speakers of English

Upon completing the questionnaire, NNSs were asked to participate in a short sociolinguistic encounter (Ash, 2002), a twenty-minute version of a sociolinguistic interview. The interviews were audio-recorded; consent for recording was requested separately and it was explained to each participant that they could still participate in the study but refuse to be audio-recorded. Every participant agreed to be audio-recorded. NSs were not interviewed.

The next part of the study was a two-part *matched-guise experiment*: 1) free profiling and 2) judgment. During the free profiling part, which always the first, the participants were asked to listen to six speakers and write anything they could about each speaker. No specific guidelines

were provided. NNSs responded on paper, NSs typed their answers. None of the participants was allowed to listen to any of the recordings more than once. During the judgment part, the participants were given a list of adjectives and asked to judge to what extent (on a five-point Likert scale) each of the speakers (same as those in the first task but presented in a different order) could be described using each of the adjectives. After this part, NNSs took a break and NSs were advised to take a break before proceeding to the next session. Those NNSs who scheduled the sessions on different days, were paid $10, for one hour of their time, after the matched-guise task.

The next session began with the reminder about voluntary consent and participant rights and proceeded with the *syntactic perception experiment*. During the experiment, participants heard 96 audio-recorded sentences and were asked to judge the naturalness of each of them on a five-point Likert scale.

The final part of the study differed for NSs and NNSs. NSs proceeded directly to the *attitudinal survey*, which included debriefing text revealing the true focus of the study. NNSs were first asked to reflect on two of the free profiles they created during the matched-guise experiment and try to remember the reasons why they described each of the speakers in a certain way, and then they participated in an *attitudinal interview*. Their consent to be audio-recorded was requested separately. At the end of the session, NNSs received their compensation.

Participants

Recruitment and Selection Criteria

Two groups of participants (native and non-native speakers of English) were recruited for the study. Native speaker participants were recruited by a call distributed via multiple course instructors who agreed to provide extra credit in exchange for participation. The students were

invited to participate in the study online, via a *Qualtrics* questionnaire (only native speaker responses were retained for analysis). An alternative assignment that required similar time commitment (approximately 60 minutes) was provided by the instructors in order to ensure voluntary participation. The distribution of the call occurred three times: at the end of the Fall semester 2017, and at the beginning and at the end of the Spring semester 2017. "End-of-semester" calls were sent only to students enrolled in language classes (Russian, French, German, Korean, etc.) or in other non-linguistic classes (e.g., Psychology), while the "beginning-of-semester" call involved students enrolled in two sections of a Sociolinguistics class (LIN 471), with the deadline for participation established early so that the students would not have been exposed to relevant sociolinguistics knowledge prior to taking part in the study.

Because participants were compensated with course credit, all survey submissions were accepted without any initial screening. Participants were simply requested to check that their sound equipment was working properly (if not, they were asked to try a different browser, check browser settings, etc.). However, the following inclusion criteria applied afterwards.

1. Were native speakers of U.S. English born in a family of native speakers or one native and one non-native speaker, as long as English was the primary language of the household throughout the participant's childhood. The participants for whom this information was not available were automatically disqualified.
2. Spoke only U.S. English as a native language (i.e., self-reported childhood bilinguals were disqualified, while advanced speakers of foreign languages learned in high school or in college were not).
3. Had not taken classes in sociolinguistics and closely related areas (e.g., "Language and Gender"). Taking basic linguistics courses such as "Introduction to language" was not

considered as a disqualifying factor, and none of the participants reported taking it. Students enrolled in LIN 471 (Sociolinguistics) were allowed to participate as long as they did so at within a week of the beginning of semester and thus were not able to acquire any sociolinguistic knowledge yet.

4. Provided basic background information and completed at least one part of the study fully.

Non-native speakers were recruited via flyers (posters) distributed across campus and via an email sent through the Office of the Registrar to all undergraduate international students enrolled at MSU at the time. The inclusion criteria were listed both on the flyers and in the email. In order to qualify, participants had to be an international student, junior or senior (i.e., be in their third of fourth year of a four-year degree), from a country where English was not widely spoken, and enrolled full-time in undergraduate classes offered at MSU. None of them reported taking classes in sociolinguistics or any related areas.

All the non-native speaker data were collected in December 2017, just before or after the end of the semester. I met with each of the non-native-speaking participants for two hour-long sessions. In most cases, the sessions were separated in time, but some participants could only meet for one two-hour block of time, which I accommodated. In the latter case, they were required to take a break between the sessions and offered refreshments (water, snacks, etc.). Each participant was paid $10 per hour, so those who participated in both sessions were paid a total of $20. Because of a misunderstanding, it was only discovered during the interview that four participants were not eligible for participation (due to either being graduate students or transfer students who had lived in the United States only for a few months). In those cases, data

collection proceeded as expected and the participants were compensated, but their data were not used for this study.

Native Speakers of English

From a total of 94 people who participated in the online questionnaire, 59 undergraduate students were retained in the study for analysis. Those excluded were either non-native speakers of English (five participants), did not complete the background questionnaire or completed only the background questionnaire but none of the other tasks. Sixty-six per cent of the retained participants were female (39 participants), thirty per cent were male (18), and the remaining two participants reported themselves as non-binary or agender. All reported growing up in the United States and speaking English as their first language, most of them in families where both parents speak English natively. All NSs were of traditional college age, between 18 and 24 (mean age = 19.56; *SD* = 1.47). Reflecting the overall demographics of the university, the majority were born and grew up in Michigan (43 NSs) or in other Midwestern states (9 NSs). Unfortunately, the information about the native speakers' race is incomplete. However, of 28 participants for whom this information was available, 27 were Caucasian.

Most participants were freshmen (21); the numbers of sophomores, juniors and senior were similar (14, 13, and 11, respectively). Nineteen reported majoring in Natural Sciences (Microbiology, Environmental Studies, Zoology, etc.), 26 majored in Social Sciences or Arts and Humanities (e.g., Political Science, Comparative Cultures, Journalism, Education, etc.), and 14 were studying various areas of business or law. Other reported majors included various areas such as Nuclear Physics or Packaging. It must be noted that the number of reported majors exceeds the total number of participants due to double-majors being reported. Only one

participant reported linguistics as an intended major but had not yet taken any classes on linguistics and could therefore be included in the sample.

Unfortunately, not all participants completed all tasks. Thus, 53 participants took the attitudinal survey (labeled “Att” in the table), 28 NSs participated in the matched-guise task, and 31 completed the syntactic judgment task (labeled “SP”). Detailed information about the participants, including their pseudonyms/nicknames and intended majors, can be found in Table 3.1.

Table 3.1
The Demographics of Native-Speaking Participants

NS*	Age	Gender	Race	Year	Intended Major**	State of Birth***	MG	SG	Att
Al	18	m	white	1	Finance	Michigan	+		
Allie	20	f	white	2	Microbiology; Russian Language	Illinois	+		+
Andrey	18	m	white	1	James Madison; Environmental Studies	Michigan	+		
April	18	f	white	1	Social Relations and Policy; Teaching Certification	Michigan	+		
anonymous	18	f	white	1	Human Biology	Michigan	+		+
b.p.	19	f	white	1	Apparel and Textile Design; Creative Advertising	Michigan	+		
bauerave	20	f	n/a	2	Environmental Studies and Sustainability	Michigan		+	+
BC393	21	f	n/a	4	Accounting	Illinois		+	+
BeyoncÃ© Pad Thai	19	f	white	3	Political Theory and Constitutional Democracy	Michigan	+		+
biggy smalls	20	m	n/a	3	Sustainable Parks, Recreation and Tourism	Maryland (Michigan)		+	+
Bunny	21	f	n/a	4	Elementary Special Education	Michigan		+	+
Candy cane	18	f	white	1	Social Relations and Policy	Pennsylvania (Michigan)	+		+
cejo	22	f	n/a	4	Nutritional Sciences	Pennsylvania (Ohio, Michigan)		+	+
Chris	21	f	black	4	Economics and World Politics	Michigan	+		+
Christian	18	m	white	1	Business	Michigan	+		+

Table 3.1 (cont'd)

crowrach	19	f	n/a	2	Biosystems Engineering	Ohio		+	+
CurveBall	19	m	white	2	Political Theory and Constitutional Democracy	Michigan	+		+
emily	19	f	white	2	Zoology; Public Policy	Ohio	+		+
Estelle	20	f	white	2	Social Relations and Policy	Michigan	+		+
Eto Chelovek	23	m	n/a	3	Nuclear Physics	Michigan		+	+
Gender Outlaw	18	nb	white	2	RCAH; Linguistics	Ohio	+		
hannad	18	f	n/a	1	Agricultural Education	Michigan		+	+
Ivan	19	m	n/a	2	Comparative Cultures and Politics	Michigan		+	+
jackal	20	m	n/a	3	Computer Science and Computational Mathematics	Michigan		+	+
Jamarcus Washington	20	m	n/a	3	Packaging	Michigan		+	+
Jannet	20	f	white	3	Marketing	Michigan	+		+
Jrob	21	m	n/a	4	Construction Management	Michigan		+	+
J-Train	22	m	n/a	4	Political Science; Pre-Law	Michigan		+	+
Kate	20	f	white	3	English; Comparative Cultures and Politics	Missouri	+		+
kaylao	18	f	n/a	1	Undecided	Michigan		+	+
kbwalsh	19	f	n/a	2	Environmental Studies and Sustainability	Illinois (Michigan)		+	+
kechosa	23	f	n/a	4	Animal Science	Michigan		+	+
kkgislason	18	f	n/a	1	Environmental Science; Management	Ohio (Missouri, Kansas)		+	+
Lena	19	f	white	1	Microbiology	Montana	+		+
Libby	21	f	n/a	4	Economics	Michigan		+	+

Table 3.1 (cont'd)

lizardgod	20	f	n/a	2	Political Theory; Comparative Cultures in Politics	Washington		+	+
longejos	20	f	n/a	3	Supply Chain Management	Michigan		+	+
lucy5	19	f	white	1	Humanities; Pre-law	Michigan	+		+
Nix	20	f	white	3	Marketing	Michigan	+		+
Patricia	18	f	white	1	International Relations and Comparative Cultures; Politics	Michigan	+		+
PUMSNC05	21	m	n/a	4	Packaging	Michigan		+	+
PW	18	f	white	1	Animal Science	Michigan	+		+
Ramirah	18	f	white	1	Arabic	Michigan	+		+
S	21	f	white	3	Political Theory	Michigan	+		+
schunkad	18	f	n/a	1	Agriculture, Food, and Natural Resources Education	Michigan		+	+
Sero	19	ag	white	2	History	Michigan	+		+
shylyk	22	f	n/a	4	Environmental Engineering	Michigan		+	+
skittles	19	m	n/a	3	Chemical Engineering	Michigan		+	+
sunshinestudios	19	f	n/a	2	Journalism	Michigan		+	+
swaswan	19	m	n/a	2	Packaging	Ohio		+	+
sydkneeb	18	f	n/a	1	Environmental Studies and Sustainability	California		+	+
Taycoy	19	f	n/a	1	Journalism; Professional Writing	Michigan		+	+
Tchaikovsky	18	m	white	1	International Relations; Finance	Michigan	+		+

Table 3.1 (cont'd)

Troy Bolton	20	f	n/a	3	Environmental Studies and Sustainability	Michigan		+	+
ttamator	22	m	n/a	4	Management	New Jersey		+	+
turtle	18	f	white	1	Social Relations and Policy; Spanish	Michigan	+		
Valerie	18	f	white	1	Comparative Cultures; Politics	Michigan	+		+
Victor Sullivan	23	m	n/a	3	Environmental Sustainability	Michigan		+	+
zdravo77	20	m	white	2	International Business Management	Michigan	+		+

* Pseudonyms/nicknames are listed as reported by participants, including capitalization. "Anonymous" was an actual nickname choice.

** Majors are listed as reported by participants.

*** If a participant reported growing up in a state different from their birth state, the second state was specified in brackets.

Note. The labels used for the last three columns represent the tasks in which each of the participants took part. "MG" stands for "matched-guise experiment", "SJ" stands for "syntactic judgment experiment", "Att" stands for "attitudinal survey".

Non-Native Speakers of English

Twenty-six non-native speakers of English (15 females, 11 males) were recruited (see Table 3.2). Each of the NNS met with me one-on-one. Eleven NNSs were juniors, 15 were seniors or recent graduates (data collection took place a few days after the end of semester). The mean age of NNSs was 22.12 (*SD* = 1.88). The length of their residence (LOR) in Michigan ranged between 17 and 84 months (mean LOR = 43.77 months; *SD* = 25.58), but for most of the NNSs the range was narrower, 29 to 41 months.

Two of the participants went to a U.S. middle school for a year, one in Michigan (Iris), one in California (Jake). The decision to retain Iris in the sample was based on the fact that, in fact, she spent a significant amount of time in an English-speaking environment (in middle school and in college) not only in the same state of Michigan but in the same area of the state (she attended school in Michigan during the time her father spent at MSU as a visiting scholar). This ensured the consistency of the vernacular elements in the input she was exposed to. As for Jake, only the matched-guise experiment data remained for him due to the audio recording failure.

One NNS participated in a short study abroad program in New Zealand, several others went to International or "American curriculum" high schools in their home country, one participant had parents who spoke English fluently and encouraged their children to learn and occasionally speak it at home. One NNS, John, arrived in the U.S. several years before entering the university, as he followed his wife who came to MSU to pursue a graduate degree, but, he spent all those years in the campus area, so the same argument of the consistency of vernacular elements in the input as I used for Iris applied to him.

Table 3.2
The Demographics of Non-Native-Speaking Participants

NNS	Gender	Age	Country of origin (L1)	LOR (months)	TOEFL scores*
Aisya	f	22	Malaysia (Malay)	29	94
An	f	20	Taiwan (Chinese)	29	90
Ange	f	22	Rwanda (Kinyarwanda)	29	106
Anna	f	20	China (Chinese)	29	93
Carlos	m	25	Dominican Republic (Spanish)	53	94
CY	m	21	China (Chinese)	41	80
Danni	f	22	Thailand (Thai)	41	
Emily	f	23	Thailand (Thai)	60	79
Frank	m	24	Pakistan (Urdu)	65	
Griggs	m	22	China (Chinese)	41	89
Ibra	m	22	The Gambia (Wolof)	29	
Iris	f	20	China (Chinese)	17	102
Jake*	m	24	South Korea (Korean)	136	88
Jenny	f	21	China (Chinese)	40	
John	m	29	Brazil (Portuguese)	77	83
Juju	f	22	Nigeria (Ebira)	41	101
Katy	f	21	Malaysia (Malay)	17	98
Lauren	f	22	Vietnam (Vietnamese)	53	
Maggie	f	21	China (Chinese)	29	80
Marsha	f	21	Indonesia (Bahasa Indonesia)	41	
Maxwell	m	21	Malaysia (Malay)	17	
MT	m	21	China (Chinese)	29	80
Ryan	m	23	Vietnam (Vietnamese)	84	97
Su	f	23	Kenya (Swahili)	29	91
Zanah	f	22	Malaysia (Malay)	29	
Zhang	m	21	China (Chinese)	53	94

Note. All pseudonyms were selected by the participants.
* Only partial data are available for the participant.

Countries of origin and L1 background of NNSs varied, with the biggest group (nine people, or 35% of the sample) speaking Chinese (Mandarin). Other L1s included Malay (four participants), Thai, Vietnamese (two), Bahasa Indonesia, Ebira, Kinyarwanda, Korean, (Brazilian) Portuguese, Spanish, Swahili, Urdu, Wolof (one).

Eighteen participants provided their TOEFL scores received prior to arrival to the U.S. (mean score = 90.88; SD = 8.18). Three participants reported IELTS scores above 7.0, also received prior to arrival. Other NNSs were unable to recall their test scores. Crucially, none of the participants received a TOEFL score lower than 79 (or its equivalent on other standardized tests), which is the minimum requirement for regular admission to MSU. As a result, none of the NNSs in this sample was required to take additional English language classes upon arrival.

CHAPTER FOUR

REMARKABLE LIKE IN NATIVE AND NON-NATIVE SPEECH: DATA ELICITATION AND RESULTS

As we saw in the literature review (Chapter Two), rLIKE as a discourse phenomenon began to attract the attention of sociolinguists in mid-80s (Butters, 1982; Schourup, 1985; Underhill, 1988), and quickly became one of the "hot topics" of variationist research due to its apparent ubiquity, functional versatility, and sociolinguistic saliency. Thirty years later, Dinkin (2016) described the amount of research done on LIKE as "enormous"; the reference list of D'Arcy's "Eight hundred years of *like*" (2017) spreads over fifty-five pages.

The study of non-native usage of rLIKE is a much more recent trend pioneered by Lee (1999) in his doctoral dissertation, in which he investigated the usage of three discourse markers (*like, I mean,* and *you know*) by Korean immigrants of three generations in the U.S. Soon after, Fuller (2003a) compared native and non-native usage of discourse markers *well, oh, y'know, like, I mean*. The first widely known and cited study of L2 usage of rLIKE, however, was Müller's (2005) book on the discourse markers (*so, well, you know,* and *like*) using the GLBCC, which included the recordings of non-native English speakers (mostly L1 German speakers) during study abroad in Long Beach (California, USA). Only a few papers investigating non-native usage of rLIKE have been published since then, many of them featuring not only rLIKE but other discourse markers of quotative verbs as well. Most of them were concerned with immigrant language learners, such as Polish teenagers in Scotland (Truesdale & Meyerhoff, 2015), beginning adult English learners in the U.S. (Hellermann & Vergun, 2007), Polish and Chinese immigrants in Ireland (Diskin, 2013, 2017; Diskin & Regan, 2015; Nestor, 2013; Nestor, Chasaide, & Regan, 2012; Nestor & Regan, 2015). Two looked at discourse marker use

by graduate students on the U.S. college campuses (Liao, 2009; Liu, 2016) and one analyzed the usage of quotatives (including *be like*) in the Mannheim Corpus of German English (Davydova & Buchstaller, 2015). The latter was one of the only two studies that targeted English speakers outside a primarily English-speaking country, the other being Algouzi's (2015) dissertation investigating the usage of discourse markers by Saudi learners of English. While Siemund, Maier, and Schweinberger (2009) used the corpora of Indian, Philippine, and East African Englishes (all countries where English is not the primary language as well), these corpora inevitably included not only L2 English speakers but also early bilinguals who grew up speaking both their local language and English, which distinguishes this study of rLIKE from other works with non-native-speaking participants. None of the studies I am aware of featured undergraduate students on an English-speaking college campus as rLIKE users. Thus, the present study makes use of a novel type of non-native participant sample.

Due to major differences in participant demographics, sample size and methods of data collection, coding and analysis, the findings of the existing studies (which I will summarize later in this chapter) cannot be easily compared, so only one claim can be made with absolute certainty: NNSs, at least those living or having lived in the target language environment, do use rLIKE in their speech. There is also evidence of NNSs being sensitive to locally-specific patterns of syntactic placement of rLIKE (Nestor, 2013; Nestor & Regan, 2015). Various social and linguistic factors may affect frequencies and patterns of rLIKE usage, but in the absence of solid evidence favoring some factors over others that would be consistent across multiple studies with participants of similar sociodemographic profiles, the best approach to any new study of non-native usage of rLIKE is exploratory rather than aimed at confirming or rejecting any pre-conceived hypotheses, and this is the approach I will take in this chapter.

This chapter will be organized as follows. In the Methodology section, I will describe the procedure I used to verify the applicability of D'Arcy's (2017) functional typology to English spoken in Michigan, then present the techniques I used to collect the usage and judgment data and methods of data treatment and analysis. Afterwards, Finally, I will present the results of the analysis of non-native patterns of rLIKE usages, as well as of native and non-native judgments of syntactic placement of rLIKE. The chapter concludes with a summary and discussion of the overall findings.

Methodology

In this section, I expand upon the methods briefly outlined in Chapter Three, to explain how I investigated non-native speakers' frequencies and patterns of rLIKE usage, as well as their sensitivity to the syntactic placement of rLIKE. Before doing that, however, I present the analysis of local corpus data which was necessary to verify that the functional typology of rLIKE developed by D'Arcy (2017) based on the data collected in Canada was fully applicable to the English spoken in Michigan.

Verifying the Functional Typology

D'Arcy's (2017) list of functions and syntactic contexts in which rLIKE may appear, as well as her claims about the frequency hierarchy among these functions and contexts (see Chapter Two), appear sound and compelling. They are based on large corpus data and supported by established developmental trajectories evident from diachronic data. However, it cannot be simply assumed that they would fully apply to speakers of a different variety of English about fifteen years later (the TEA corpus data which D'Arcy used were collected in 2002-2006).

In order to establish a contemporary reference point for the twenty-six NNSs I interviewed for this study (see Chapter Three for details about the participants), I extracted and

analyzed all tokens of LIKE (a total of 401) from a small subsample of the IHELP-MI corpus[8]. The subsample comprised interviews with ten white, college-aged speakers (seven women and three men), all Michigan-born and raised. I exhaustively extracted LIKE tokens starting from the second ten minutes of each interview. This is not the procedure D'Arcy used in her study of the TEA archive of Toronto English. Recall that D'Arcy extracted all potential syntactic/functional contexts for rLIKE and then recorded the presence vs. non-presence of LIKE in those contexts. This provided a measure of frequency of rLIKE in each context. In contrast, and for convenience, my approach was variant-centered rather than variable (i.e. context)-centered. This is also a standard approach in research on discourse pragmatic variation (Dinkin, 2016) (D'Arcy, 2017). As a result, I cannot directly compare the frequencies reported by D'Arcy with the frequencies observed in my data. But I was able to confirm that all the major functions of rLIKE characteristic of Canadian English (approximator, quotative, discourse marker, discourse particle) also appear in the speech of Michigan college youth. I then looked closely at the syntactic contexts in which mLIKE and pLIKE appeared, to further provide some support for the assumption that the distribution of mLIKE and pLIKE is not substantially different from that observed by D'Arcy for Canadian English.

As we have seen, D'Arcy (2017) reports that within the clausal domain, matrix CP is the oldest and thus the most frequent context for mLIKE, followed by subordinate CP. Subordinate TP is the newer context, so it may not be a part of repertoire of older speakers at all and among the younger speakers it would be the context in which mLIKE least frequently appears. The same frequency hierarchy of matrix CP > subordinate CP > subordinate TP was also observed in

[8] IHELP-MI = "The Influence of Higher Education on Local Phonology," a 2013-2016 project conducted at Michigan State University that resulted in a corpus of recorded and transcribed interviews with native-speaking MSU students (Wagner, Mason, Nesbitt, Pevan, & Savage, 2016).

IHELP-MI. Out of 65 mLIKE tokens, the overwhelming majority (54) were in the matrix CP contexts. Ten tokens were in subordinate CP contexts, and only one token of mLIKE in subordinate TP context was found. We can conclude that Michigan English speakers evince the same syntactic pattern for mLIKE as Canadian English speakers. Thus, in this chapter, it will be possible to make direct comparisons between D'Arcy's more extensive study of mLIKE and my own findings for non-native speakers in Michigan.

As far as pLIKE is concerned, we can draw the strongest conclusions for the nominal domain. As we saw above, D'Arcy finds that the DP context is one of the oldest contexts, and thus one of the most frequent hosts for pLIKE. The nP context in contrast is recognized by D'Arcy (2017: 137) as a brand-new addition to the pLIKE context repertoire, being restricted to speakers born after 1970. In IHELP-MI, the same ranking order of DP > nP was observed: pLIKE in the DP context predictably accounts for 40.9% of all pLIKE tokens (45 out of 110), while pLIKE in the nP context accounts for 8.7% of all pLIKE tokens (9/110). However, nP is the second most frequent adjunction site for pLIKE in the subset of IHELP-MI I analyzed, making it more frequent than developmentally older contexts in other domains (verbal and adjectival). It is also important to note that these nine tokens were produced by five different speakers (four women and one man) and thus nP should be recognized as an established context for English spoken by Michigan youth. We can therefore conclude that Michigan speakers observe the same hierarchical frequency relationship between DP and nP contexts as in Toronto, but overall the nP context is unexpectedly high ranked.

Indeed, the AP context is only the third-ranked context for pLIKE in IHELP-MI (eight tokens produced by five different speakers, 7.2% of all pLIKE tokens). However, all of these tokens feature rLIKE in predicative structures (e.g., *Um, rock is LIKE heavier with usually real*

instruments, you know, playing (IHM1-45, f) and there is not enough of them to compare the patterns of adverb placement recognized by D'Arcy (preceding or following LIKE).

Within the verbal domain, differences may be expected between "significantly more frequent" (p. 149) auxiliary and infinitival *to* contexts and less frequent contexts with (semi-) modals and bare finite verbs. However, low token numbers make comparison meaningless: in IHELP-MI, pLIKE appears seven times in the infinitival *to* context, a total of seven times in both modal and non-modal auxiliary contexts, and only three times on the left of the bare main verb. It is also notable that, even though D'Arcy stated that pLIKE occurs only to the immediate left of the lexical verb, in contemporary American English spoken in Michigan this is likely to be a preference rather than a categorical constraint. In IHELP-MI, I found two tokens of pLIKE used before *to:* for example, *It feels really big, like I've always been up here for like sporting events and LIKE to visit people that have been up here* (IHM-22, f).

Overall, it is evident that D'Arcy's findings for Canadian English speakers' syntactic distribution of mLIKE and pLIKE are not markedly different from the patterns observed in Michigan English, even though only a small number of tokens was analyzed. The level of functional and syntactic diversity in non-native rLIKE use may serve as a good measure of acquisition of rLIKE as a discourse element. In sum, analysis of the IHELP-MI subsample confirms that for the rest of this chapter, comparisons may be reliably made between non-native speakers of English in Michigan and, where necessary, D'Arcy's large-scale study of native speakers of English in Canada. Wherever possible, I will also provide direct comparisons between NS and NNS usage in Michigan.

Data Elicitation

In this section, I discuss the actual elicitation measures I used, a sociolinguistic encounter, and a syntactic judgment experiment.

Short Sociolinguistic Encounter

The short sociolinguistic encounter, a twenty-minute version of a sociolinguistic interview developed by Ash (2002), was two-fold in purpose. First, the information about the participants, received through the background questionnaire, could be confirmed (triangulated) and extended. However, the main purpose of the interview was to elicit rLIKE in non-native speech. While I had a set of prepared questions (see Appendix C), I did not follow the protocol fully with each of the participants, which is in line with sociolinguistic practices (see, for example, Labov (1984)). After establishing initial trust, I followed the participants' lead and pursued various potentially emotion-provoking themes. Some of the participants preferred to discuss the differences between their home countries and the United States, others talked about learning to cook food from their home country as only American food was available in local canteens and cafes, some discussed their relationships with their parents (e.g., fights about career choices), etc.

Syntactic Judgment Experiment

The syntactic judgment task is a listening experiment in which participants heard 96 sentences (see Appendix D) and judged their naturalness. The exact instruction given to the participants was the following: "You will listen to a number of sentences. For each one, please indicate how natural the sentence sounds to you on a scale from 1 (unnatural) to 5 (perfectly natural). In other words, decide whether the sentence sounds like something that you might hear in everyday casual speech, for example, on campus. It's fine if you do not understand the broader context of a given sentence."

Most of the sentences presented in the experiment were selected from a subset of IHELP-MI, that is, they were produced by college-aged native speakers of English. The selection of the stimuli was conducted in several steps. First, a subset of ten IHELP-MI speakers (seven females, three males) was selected. Then, all tokens of LIKE were extracted from the second ten minutes of speech produced by each speaker in their sociolinguistic interview. They were coded by me and one independent coder, following D'Arcy's (2017) classification. The first level of coding distinguished between the unremarkable and remarkable functions of LIKE; a separate code "other" allowed for separating ambiguous tokens. After performing the reliability test on 413 tokens (Cohen's Kappa = 0.67), only the undisputed tokens of LIKE were retained. Out of those, the list of 22 stimuli (target items) was created. The items included in the list illustrate three of the remarkable functions of LIKE that are sociolinguistically salient (Maddeaux & Dinkin, 2017): the mLIKE (clause-initial discourse marker), the pLIKE (clause-medial discourse particle), and aLIKE (approximator). Ten additional stimuli representing non-local, constrained or rare usages of rLIKE were taken from D'Arcy (2017), with the author's consent. All sentences were audio-recorded by a volunteer actress, a young woman born and raised in Michigan.

The first two functions were represented in different syntactic positions, so the complete list of stimuli included the contexts listed below. Items taken from D'Arcy (2017) are marked with an asterisk.

1. Approximator rLIKE (two items)

    (a) *My sister's LIKE twenty-seven.*

2. Discourse marker rLIKE (eight items, two per context)

    2.1. matrix CP, sentence-initial (b) or non-sentence-initial (c)

    (b) *LIKE I knew I wanted to go to Michigan State.*

(c) *It was considered on campus housing, and LIKE everything was through an agency.*

2.2. subordinate CP (d) and subordinate TP (e)

(d) *I mapped out LIKE where I have to go for my classes yesterday.*

(e) *One of my cats meows so much 'cause LIKE he's really picky and everything.**

3. Discourse particle rLIKE (sixteen items; two per context)

3.1. Nominal domain: DP, functioning as an argument of vP (f) or a complement of PP (g); nP (h)

(f) *And um, the fraternity guys are getting in groups of ten and doing LIKE challenges.*

(g) *I'm not so much into LIKE European History and that type of stuff.*

(h) *So, um, my mom, she served on a, um an agriculture LIKE advisory board.*

3.2. Verbal domain: before to-infinitive (i); within to-infinitive (k); to the left of main verb (l); between modal and main verbs (m)

(i) *I've always been up here for sporting events, and LIKE to visit people that have been up here.*

(k) *My goal isn't to make money… My goal is to LIKE make a difference in people's lives.*

(l) *And then you go through dungeons and LIKE fight monsters.*

(m) *But I also like that I can LIKE stay here for school.*

4. Sentence adverb (two items), (n)

(n) *We need to smarten it up a bit LIKE. **

5. Constrained usages: before a copula (o-p); before a personal pronoun (r)

(o) **He LIKE was so happy to take a bath. **

(p) **I've caught trout that LIKE are small. **

(r) ***I had a crush on LIKE him. **

6. Authentic rare usages: within a lexeme (s); before an idiom (t); within a long verb phrase (u)

(s) *She's very aware of her feelings but is un-LIKE-sympathetic to others.**

(t) *We all are LIKE down to earth type of people. **

(u) *But I've never actually LIKE been up here. **

Besides the target stimuli, 32 sentences not containing rLIKE were randomly extracted from the same subcorpus of IHELP from which the majority of stimuli were extracted. These sentences served as authentic fillers. Some of these sentences contained other discourse markers (e.g., "well", "and", "I mean") or other elements of the vernacular, such as clause-final "so" or general extenders ("and stuff", "or whatever").

Finally, three other types of sentences, all constructed, were added (see Appendix D). The first type represented ungrammatical sentences (a total of 22), constructed so that they would contain grammatical errors commonly occurring in non-native speech (e.g., erroneous subject-verb agreement or negation, pluralized uncountable nouns, misplaced adverbs, to-infinitive following modal verbs, etc.). The second type (seven sentences) included elements of various stigmatized vernaculars (e.g., African American Vernacular English and other regional dialects; the elements include double negation, double modals, etc.). Finally, sentences of the

third type (a total of three) contained inverted idiomatic or common expressions, such as “white and black TV”, “cheese’n’mac”, and “dirty and quick”.

They were presented (in random order, unique for each participant) to NNSs via *PsychoPy* and to NSs via *Qualtrics*; both groups of participants responded by a mouse click on an appropriate dot on a 1-5 Likert scale.

Data Analysis

Short Sociolinguistic Encounter. After the data collection was completed, all interviews conducted with NNSs were carefully transcribed by me or by undergraduate research assistants who received prior training. Express Scribe Transcription Software was used for transcription.

All tokens of LIKE were extracted from the transcriptions using AntConc software (Anthony, 2017) and then coded as remarkable, unremarkable, or ambiguous. Then, all remarkable tokens of LIKE were further coded for function/syntactic position (D’Arcy, 2017), as discourse markers, discourse particles, quotatives, approximators, or ambiguous tokens. Ambiguous tokens were subsequently discarded. The initial attempt to code discourse particles further was not entirely successful due to the difficulty of analyzing the syntax of non-native speech. Thus, it is difficult to decide whether the context in which pLIKE appears in the following example can be categorized as the nP context: *I don’t have a specific LIKE examples in my mind right now.* On the one hand, LIKE is placed between a determiner (in this case, indefinite article) and a noun, but on the other hand, the indefinite article should not have been used in this phrase at all.

Thus, only frequencies of “larger” rLIKE functions (marker LIKE, particle LIKE, quotative LIKE, approximator LIKE) were calculated and subsequently normalized as

frequencies per 1000 words. No inferential statistic methods were used for the analysis of rLIKE usage; descriptive statistics and visual (graphic) representations were used instead.

Syntactic Judgment Experiment. The results of the *syntactic judgment experiment* were analyzed using both descriptive and inferential statistics (a series of independent-sample *t*-tests). In addition to that, color-coding and visual (graphic) representations were used to reveal response patterns across the entire sample.

Results

In this part of the chapter, I will first present the distribution of rLIKE by function in native and non-native speech. Then I will narrow the focus of the study and explore the individual variation in the patterns of rLIKE usage observed in non-native speech, as well as social factors contributing to this variation. Finally, I will present the results of syntactic judgment experiment and discuss the extent to which NNSs are sensitive to native patterns of syntactic placement of rLIKE.

Distribution of rLIKE across Functions in Native and Non-Native Speech

In order to find whether NNS have acquired rLIKE at all, and if so, whether they have acquired it at native-like rates across remarkable functions, I extracted a total 963 tokens of LIKE from the interviews (short sociolinguistic encounters) with 26 NNSs. 774 tokens (80.4%) were those of rLIKE. This proportion of rLIKE to all tokens of LIKE was even higher than the one found in NS data in the IHELP-MI subsample (71%, 286/401), although this may be due to difference in the size of the data sets. Especially important are the proportions of different functions of rLIKE which were strikingly similar across native and non-native speakers (see Table 4.1).

Table 4.1
Distribution of rLIKE Tokens across Functions in Native and Non-Native Speech

Function of rLIKE	NSs		NNSs	
	N	%	N	%
(Overall)	286	100%	774	100%
particle	110	38.5%	347	44.8%
marker	65	22.7%	196	25.3%
quotative	44	15.4%	79	10.2%
approximator	25	8.7%	40	5.2%
ambiguous	42	14.7%	112	14.5%

Note. The "ambiguous" category includes tokens that were difficult to classify (e.g., as an approximator or a discourse particle), as well as false start tokens and tokens found between clauses and surrounded by pauses in speech.

As can be seen in the table, the order of frequency is the same for both native and non-native speakers: clause-medial pLIKE is the most frequent, followed by mLIKE, then the quotative, and finally the approximator. The only notable difference is between the proportion of quotative LIKEs used (NSs use more), but without additional analysis, it is not possible to tell whether this difference could be attributed to the differences in narrative styles between native and non-native speakers (e.g., NNSs might be less likely to produce narratives at all or quote people's speech while narrating events) or to the different preferences in quotative use (e.g., NNSs might prefer a different quotative verb, such as *say*, over LIKE, while LIKE may be the quotative verb of NSs' choice).

Discourse Marker (mLIKE). Within the clausal domain, the similarities remain (see Table 4.2). The proportion of mLIKE tokens in the most frequent matrix CP context is comparable in native- and non-native data (83.1% and 73.8%, respectively), the same applies to the subordinate CP context. Curiously, 14 tokens of mLIKE at subordinate TP were found in NNS data (7%), while only one such token was identified in the subcorpus of IHELP-MI. Eight

different NNSs produced the TP tokens, so such usage was not restricted to any single native-like outlier in the group.

Table 4.2
Distribution of mLIKE Tokens across Functions in Native and Non-Native Speech

Function of mLIKE	NSs		NNSs	
	N	%	N	%
(Overall)	65	100%	195	100%
mx CP	54	83.1%	144	73.8%
sub CP	10	15.4%	37	19.0%
TP	1	1.5%	14	7.2%

However, 11 out of these fourteen tokens were those of LIKE following the subordinator "because", or "cause", as in 5a and 5b.

(5) a. *Like sometimes, some of the patients like don't get their needs met because LIKE there's often also like… also a delay in like for the medical director to sign those forms…* (Juju, f, 22, Nigeria)

b. *The first, I don't think like they are very good chess because LIKE I got a lot of medals during my high school, during my last few years…* (Maxwell, m, 21, Malaysia).

Notably, the only example from IHELP also features "because like" (twice, in fact, even though the second time the adjunction site is not at TP): *Cause LIKE people went to Grand Valley cause like nursing* (IHM-36, f). Discussing the distributions in TEA, D'Arcy (2017) notes that "*because like* is fairly robust relative to this sector (TP adjunction)" (p. 92). Based on the fact that its frequency stabilizes at the 8% level for speakers born after 1973, even though overall frequency increases, she concludes that "*because like* does not represent insipient fixation"

(ibid.), which means that it cannot be treated as a new discourse marker that should be studied separately from LIKE.

This conclusion, however, cannot be simply applied to NNSs, so there is still a possibility of chunked, or, in D'Arcy's terms, "routinized" usage. While the low number of tokens does not allow for any in-depth analysis, the following example (6) may serve as circumstantial evidence in favor of the routinization hypothesis:

(6) *I think it's a really good way of teaching cause LIKE for example we were on the glacier and the professor talks LIKE how the glacier forms, LIKE what are the parts called, like those kinds of things* (Iris, f, 20).

In this example, the speaker used the discourse marker LIKE three times. The two latter usages both feature LIKE in the subordinate CP contexts (LIKE precedes the subordinator) but with different and non-causative subordinators (*how* and *what*). The former context, however, is that of subordinate TP with (be)*cause* as subordinator. Given that all three usages were produced by the same speaker and, furthermore, within the same sentence, it is possible to speculate that it is the subordinator that motivates the placement of LIKE.

Discourse Particle (pLIKE). The similarities between native and non-native distributions are also evident with regard to the clause-medial pLIKE (see Table 4.3). In NS data, 40.9% of all pLIKE tokens were found in DP contexts; in NNS data, this proportion was 46.4%. As for the newer nP context, the same number of pLIKE tokens (nine) was found in both data sets, but with the difference in corpus size, it constitutes 8.2% of all NSs' pLIKEs and only 2.6% of NNSs'. This difference is to be expected, given that the infrequency of the context in native speech means it is infrequent in the input NNSs receive.

Table 4.3
Distribution of rLIKE Tokens across Contexts in Native and Non-Native Speech

Contexts of pLIKE	NSs		NNSs	
	N	%	N	%
(Overall)	110	100%	347	100%
DP	45	40.9%	159	45.8%
nP	9	8.2%	9	2.6%
between auxiliary and main verb	7	6.4%	17	4.9%
within *to*-infinitive	7	6.4%	15	4.3%
PP	4	3.6%	23	6.6%
left of bare main verb	3	2.7%	31	8.9%
other/ambiguous tokens	35	31.8%	93	26.8%

Note. The category "Other/ambiguous tokens" included infrequent contexts (e.g., pLIKE within a passive construction), clearly clause-medial false starts, and tokens I was unable to classify with enough certainty.

The only notable difference is within the verbal domain: NNSs use pLIKE on the left periphery of the bare main verb more often than NSs. According to D'Arcy (2017), it is one of the least frequent contexts for pLIKE, which is reflected in the extremely low token number in the native data. NNSs, however, used almost 9% of their pLIKEs in that context. One possible speculation is that NNSs may use more bare main verbs overall.

To summarize, the NNSs who participated in my study were remarkably native-like with regard to the syntactic positions in which they use rLIKE: Not only do they use it in a wide range of positions, but the distribution of all rLIKE tokens across these positions matches the distribution observed in the native speaker data. Crucially, as both data sets were collected on the same college campus with only a few years of difference between the times of collection, it can be argued that the native data accurately represent the input these particular NNSs were exposed to during their residence in Michigan. As most of the NNSs reported learning about rLIKE as a discourse phenomenon only upon arrival in Michigan, which was also their first experience of living in an English-speaking environment, it indicates that they were able to acquire rLIKE and incorporate it into their repertoire within about two years.

Non-Native Usage of rLIKE: Individual Variation and Factors Affecting Usage

In the previous section, I presented evidence to show that NNSs collectively demonstrated a high degree of native-likeness in their syntactic placement of rLIKE across various rLIKE functions (following D'Arcy's (2017) typology). However, the degree of inter-speaker variation in usage was high, which, although expected based on existing SLA literature (e.g., Hellermann & Vergun, 2007; Liu, 2016; Müller, 2005), requires detailed analysis and interpretation.

None of the twenty-five NNSs who were interviewed for this study failed to produce at least one token of rLIKE. Indeed, three top rLIKE users (Danni, Juju, and Zhang) produced rLIKE at a rate of over 75 tokens per 1,000 words (Figure 4.1). In contrast, four participants -- MT, Jenny, An, and John-- produced a total of fewer than ten tokens each, giving normalized rates of 7.64, 5.69, 2.7, and 2.22tokens per 1,000 words, respectively.

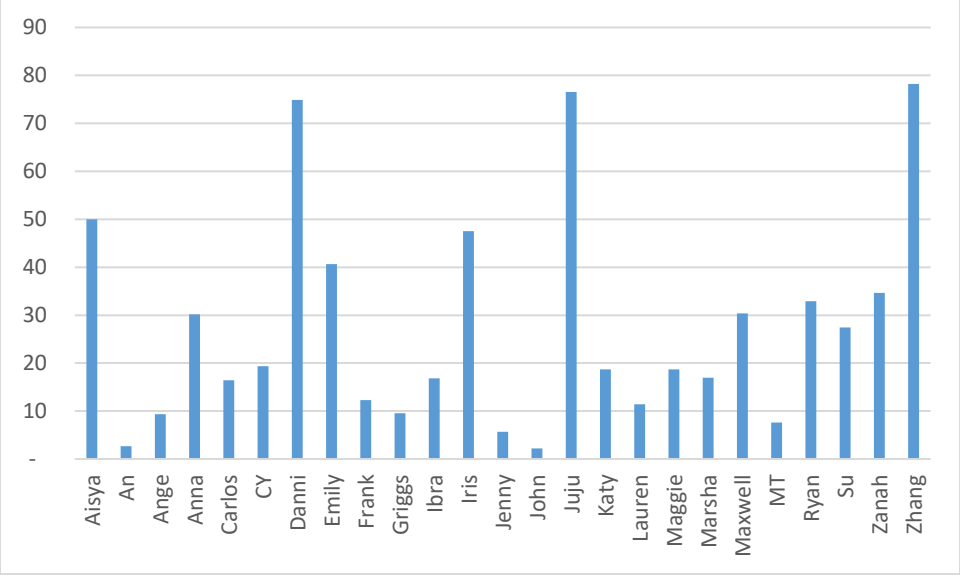

Figure 4.1
Inter-speaker Variation in rLIKE Usage by NNSs (Normalized per 1,000 Words). The *y*-axis represents normalized frequency of rLIKE usage per 1,000 words, the *x*-axis represents the study participants.

Nine speakers used the full range of rLIKE functions: approximator, quotative, clause-initial marker, and clause-medial discourse particle; another seven speakers used three of the four functions (missing either a quotative of an approximator, the most content-dependent functions). Therefore, it can be argued that 60% of the participants have the full range of rLIKE functions in their repertoire. Furthermore, three of the four non-users (those who produced fewer than 10 tokens) used their rLIKEs in more than one function.

Since the group is relatively homogenous with regard to age (while the overall age range is 20 to 29, only three of the participants were 24 or older), age cannot be considered as a factor influencing rLIKE usage among the participants of this study. Age of arrival cannot be considered either, since all participants entered the US for the first time as college students. Instead, I focused on speaker gender, region of origin, length of residence, friendship network, and intention to stay in the U.S. after graduation as a proxy for the potential level of investment into their L2 identity (a measure inspired by Diskin and Regan's (2015) findings on the relationship between the type of migration and usage of the symbolically Irish clause-final LIKE). Finally, I will discuss the relationship between beliefs about and attitudes towards LIKE and the usage of LIKE.

Speaker Gender. Among the 25 participants, 15 were female and 10 were male. If the four non-users are excluded from analysis (which was done for all subsequent calculations), the proportion stays the same, with 13 female and 8 male speakers. Two of the three top users (rate of 70+ tokens per 1,000 words) are female, but the most frequent user, Zhang, is a man. Overall, the women appear to have been more successful at acquiring rLIKE (see Figure 4.2), using rLIKE at a rate of 36.2 tokens per 1,000 words (vs. 23.5 among men).

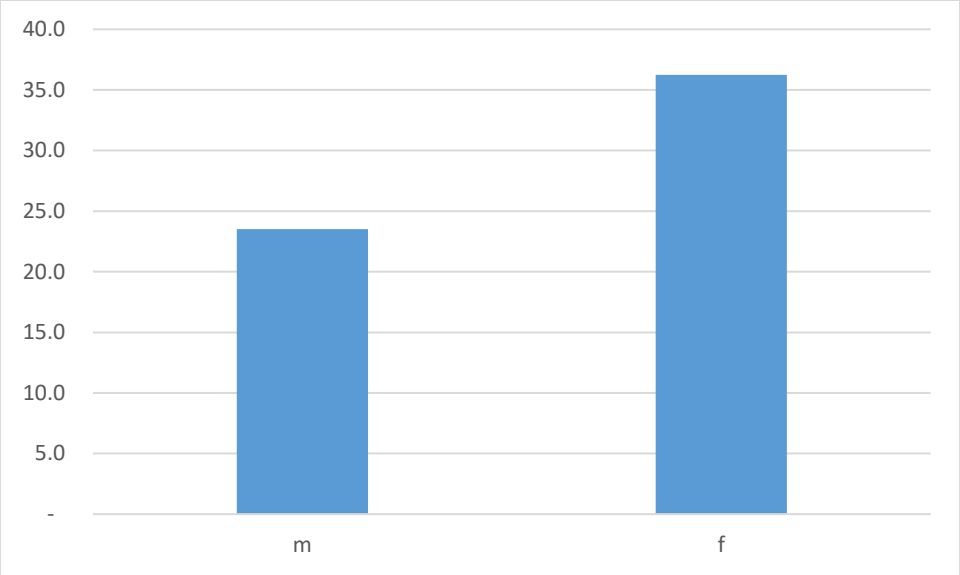

Figure 4.2

Frequencies of rLIKE Usage per 1,000 Words by Male and Female NNSs (n=21). The *y*-axis represents normalized frequency of rLIKE usage per 1,000 words, the *x*-axis represents the participant gender (m = male, f = female).

Women NNS also seem to have been more successful than men in the acquisition of two rLIKE functions, namely marker and particle LIKE. Women use clause-initial marker LIKE at a rate of 12 tokens per 1,000 words, while men use it at a lower rate of 8. More pronounced differences exist with regard to clause-medial particle LIKE: While men use it at almost the same rate as markers (9 tokens per 1,000 words), for women the rate is almost twice as high, 17 tokens per 1,000 words, which means that it is the difference in the use of that particle that mainly contributes to the gender division in the overall rLIKE frequencies. Approximator LIKE appeared to be a gender-neutral function, as men and women used it at very similar and very low rates of 1.3 and 1.5 tokens per 1,000 words, respectively. The situation with the quotative was similar (rates of 3.2 for men and 2.9 for women).

Region of Origin. While existing literature on non-native usage of rLIKE does not indicate that first language background or country of origin may be contributing factors in it, I considered it important to investigate this possibility. It is known that partial equivalents of

rLIKE exist in other languages, such as German (see Golato (2000) on quotatives *und ich so/und er so*), French (see Fleischman and Yaguello (2004) on discourse marker *genre*), and Spanish (see Kern (2014) on discourse marker *como*). Anecdotal evidence from my own data set suggests that such equivalents may be also present in Swahili, Thai, and Japanese (the participants mentioned those during attitudinal interview). Unfortunately, no systematic cross-language studies on rLIKE equivalents have been published yet, which makes it impossible to make testable predictions with regard to all L1s in my data but may suggest a route of investigation if speakers from a certain L1 background demonstrate higher than average rates of rLIKE usage. However, as many of the participants were unique representatives of their respective countries in this study, it was impossible to consider the country of origin as a factor influencing rLIKE usage. However, I chose to divide the participants into four groups. The first group included nine participants from Southeast Asia (Malaysia, Thailand, Vietnam, Indonesia), the second group consisted of speakers from China and Taiwan (six NNSs), the third included four NNSs from African countries (Rwanda, The Gambia, Nigeria, Kenya), and the fourth was, in essence, not a group but a "miscellaneous" category and consisted of just two speakers, one from Pakistan and the other from the Dominican Republic. This division, while not random, can be argued to be meaningless, because only the members of the China and Taiwan group share the same L1 (in my sample, all participants from these countries were native speakers of Mandarin Chinese). However, there is a common agreement among anthropologists and sociologists that people from Southeast Asian countries share certain cultural practices and can be grouped together, especially when studied outside of their home countries, in immigrant settings (e.g., see Liamputtong (2006); Reyes (2007)). It can be speculated, therefore, that there can be commonalities at the level of discourse (e.g., practices of narration, which may involve preferences for quoting direct

speech or retelling, thus influencing the usage of quotatives, etc.), which, in turn, could influence the usage of such discourse elements as LIKE. From this perspective, it was interesting to compare speakers from these regions in terms of their rLIKE usage.

As can be seen in Figure 4.3, speakers from Southeast Asia appear to be the most frequent users of rLIKE (the rate of 36.3 tokens per 1,000 words); the speakers from China and Taiwan and from African countries show slightly lower and very similar rates of rLIKE (31.3 and 29.9, respectively). All three groups use the marker LIKE at almost the same rates (11 for Southeast Asia and China and Taiwan, 12 for Africa), the rates of approximator use are very low (1.5, 2.3, and 0.7). With regard to the particle, Southeast Asians lead (rate of 17), while Chinese and African speakers demonstrate the same rate of 13 particles per 1,000 words. In quotative usage, Southeast Asians also lead (4.4 tokens per 1,000 words), followed by speakers from African and "other" countries (the rate of 2.8 in each group). Speakers from China and Taiwan, however, demonstrated a very low rate of 1 quotative token of rLIKE per 1,000 words. I did not calculate the number of quotes each participant used, and therefore cannot definitively state that the low rate of qLIKE in the speech of L1 Chinese participant was a result of avoiding direct speech quotation rather than of a preference towards a different quotative (e.g., *say*). However, I would like to speculate that this could be the case, based on the impression I formed while interviewing speakers and analyzing interview transcripts.

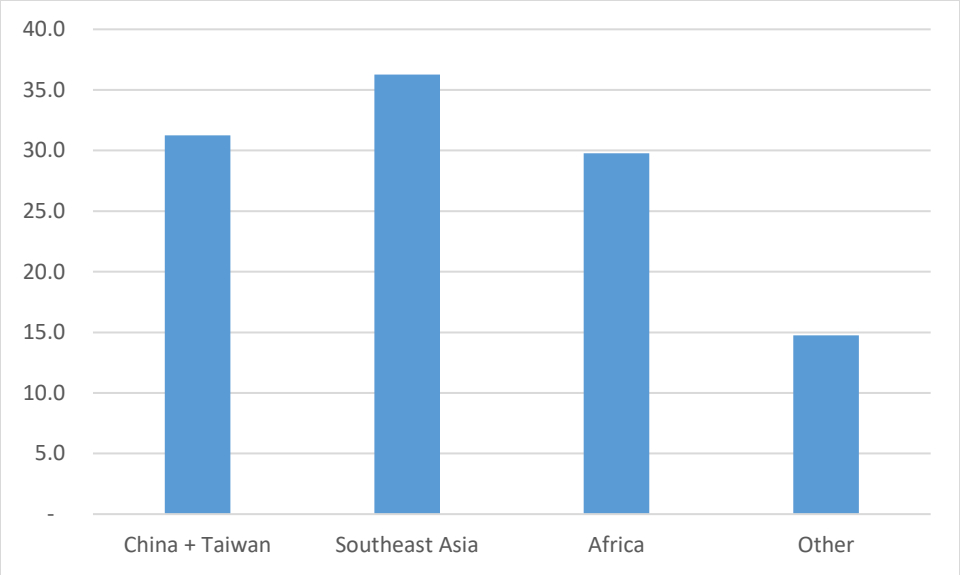

Figure 4.3
Frequencies of rLIKE Usage per 1,000 Words by NNSs from Different Regions of Origin

While it is technically possible to compare the usage of rLIKE by men and women from each region, the results would be not be particularly meaningful due to small sample size (e.g., there is only one man among speakers from Africa). The overall trend of females demonstrating higher frequencies, however, holds for rLIKE usage among both Chinese and Southeast Asian speakers (see Figure 4.4).

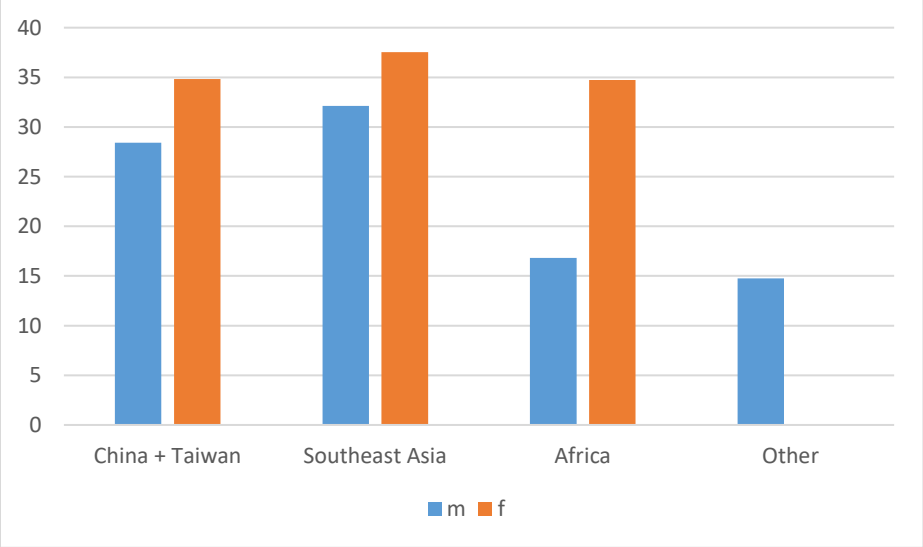

Figure 4.4
Frequencies of rLIKE Usage per 1,000 Words by Male and Female NNSs from Different Regions of Origin

Length of residence. The shortest reported length of residence among the NNSs in this study was 17 months (three participants), the longest (77 and 84 months) were reported by two participants who immigrated to the United States before beginning their college studies. The decision to include these participants in the final sample was based on the fact that they had never been to another English-speaking country prior to immigration and had lived in the campus area during their entire stay in the country, which ensured that the variety of English they had been exposed to was the same as for the other NNSs in this study.

The most frequently reported LOR was 29 months (nine participants), and the mean LOR was slightly over 40 months. No apparent correlation was found between LOR and the frequency of rLIKE usage, both in general and by function ($r = 0.04$). This lack of correlation can be illustrated by reporting LOR of the top five and bottom five rLIKE users. For the top five, their respective LORs were 53, 41, 22, 29, and 17 months, while for the bottom five it was 29, 29, 40, 29, and 77 months. In other words, LOR could not be described as an important factor affecting the frequency of rLIKE usage in this study, although it must be noted that this lack of correlation could be, in part, a result of using an LOR threshold as a participant selection criteria for this study, so that no new arrivals with no exposure to the native speech on campus were included.

Friendship network. Although I did not use any specific measures estimating the amount of exposure to native English speech (e.g., journals or time logs), I decided that an indirect measure of reported friendship network could be employed. Instead of applying any developed protocol of social network analysis (e.g., Milroy, 2000), I simply coded the responses received during the interviews using the following categories: (0) the person socialized mostly with people from their home country who speak the same L1, (1) the person socialized mostly with other international students but not from the same country, so English would be used as a

lingua franca, (2) the person has a diverse circle of friends that consists of other international students and Americans (native English speakers), (3) the person socialized mostly with Americans (native English speakers).

Out of 23 NNSs for whom these data were available, four people belonged to Group 0, four to Group 1, Group 2 was the most populated (13 participants), and only two speakers reported socializing mostly with Americans (Group 3). As with LOR, no apparent connection between the friendship network and rLIKE usage was found. Thus, out of Group 3 speakers, one was a non-user (rate of 2.7 rLIKEs per 1,000 words) and the other user rLIKE at a relatively low (within this sample) rate of 9.55 tokens per 1,000 words. Group 0 speakers, on the other hand, were evenly distributed along the frequency scale and included the top user, Zhang, who used 78.25 rLIKEs per 1,000 words. Similar distributions applied to members of Group 1 and Group 2. It is clear that the type of friendship network, at least when measured by relying on participant self-reports, does not help shed any light on the differences in the frequency of rLIKE usage.

Intention to stay in the U.S. As Diskin and Regan (2015) found a connection between rLIKE usage and the level of investment into L2 identity expressed as the reasons for being in and intention to stay in the country in their Irish data, I decided to explore a similar possibility in my data as well. Of course, the participant responses to the question about their future plans and intentions depended on the stage of their studies at which they were at the time of the interview. While some of the participants were only in the middle of their second year of studies and their responses were mostly the preferences they expressed, others were fresh graduates (the graduation ceremony took place a few days before the interviews) who already had concrete plans for their future. For some others, choice or preferences were not even involved, as they

were the recipients of scholarships from the governments of their countries and were required to return home after graduation.

Nevertheless, I was able to divide the participants into four groups based on their responses: (0) the person intends to return to their home country upon graduation; (1) the person would prefer to stay and work in the U.S. for a while to gain valuable experience but they would prefer to later return to their home countries because they could not see themselves living in the U.S. for a long time due to cultural differences and/or separation from family[9]; 2) the person would like to neither go home, not stay in the U.S., living in a third country would be the preferred option; 3) the person would like to stay in the U.S.

Out of 23 NNSs for whom these data were available, six people intended to return to their home country immediately upon graduation (including four who were required to do so), while three speakers expressed a desire to first explore job opportunities in the U.S. and possibly other countries but to return home when they become older. Seven NNSs indicated a preference for living in a third country. Their proposed motivations varied greatly, from salaries in a certain field being the highest in a certain country, to a desire to provide humanitarian aid in remote areas of third-world countries, to wanting to travel as much as possible. Another seven participants expressed a firm desire to stay in the United States, although motivations also varied from value-based to more practical. For example, one participant said that U.S. was a more gay-friendly country than their homeland, while another wanted to make a career in the movie industry of which the U.S. is the world capital. Among these seven who expressed a desire to stay in the U.S., four NNSs were frequent rLIKE users (with rates over 30 tokens per 1,000

[9] For example, Ange (f, 22, Rwanda) said: "I have no preferences so far, cause I feel like where the opportunities is I will just go there. I don't have any specific place that I want to stay in, but I know for sure that I want to stay in my country when I turn thirty-forty, something like that. … Why, because, cause I just feel like it's just good to be in my country as I age, cause it's where I feel love, it's where I feel love and where I feel happy."

words; the group included the top user), and two were non-users. As for the other two groups, their members were evenly distributed along the frequency scale, thus providing no reason to treat the intention to stay as an important factor affecting rLIKE usage (see Table 4.5 later in the chapter).

Beliefs about and attitudes towards LIKE. As I stated in the literature review, there is evidence in recent variationist research that NNSs may associate socially salient vernacular variants (e.g., *t*-glottaling in British English) with "native-speakerness" and thus overuse it in any situation when they would want to signal their native-likeness and to present the best version of their L2-speaking self (Cheshire et al., 2011; Kerswill, 1994; Modaressi-Tehrani, 1978; Sharma, forthcoming).

Given this evidence and the fact that rLIKE, indeed, is a highly salient element of American English vernacular, I decided to check if the frequency of usage corresponds with the participants attitudes towards LIKE and their recognition of it as a specifically American feature. It turned out to be a promising line of inquiry.

First, I analyzed the attitudinal interviews (see Methodology) of the top five and bottom five users with respect to their normalized frequency of rLIKE usage. I coded overall attitudes towards rLIKE as negative (e.g., "Like is a marker of poor English"), positive (e.g., "It makes speech more understandable"), and neutral (e.g., "Everyone speaks this way"). Among the top five users, three demonstrated neutral attitudes towards it, one speaker showed positive attitude, but the most frequent user, Zhang, called it a marker of poor English and said that people should avoid using it. Among the bottom five users, three expressed their dislike for LIKE and two had no positive or negative feelings towards it.

However, a more revealing finding was that all five top users, regardless of their attitudes towards LIKE, positively identified it as a specifically American phenomenon (e.g., calling it "100% American" or "an indicator of people having lived in the U.S for a while"). Conversely, only one of the bottom five users noted that "Americans use it a lot", with the other one identifying a particular group of Americans ("American girls") as rLIKE users. When I mapped the participant responses to the question about what people, in their opinion, could be described as frequent LIKE users, a clear trend emerged: The most frequent users (rates of 40 tokens per 1,000 words and higher) were unanimous in their recognition of LIKE as an American phenomenon. With less frequent users, there was more variability, as many of them associated LIKE not with Americans in general but with a certain group of Americans (white Americans, black Americans, American teenagers, etc.), and some users with medium rates did not associate LIKE with Americans. However, this evidence allows me to speculate that the association between LIKE and Americanness is at least a contributing factor to the frequency of LIKE usage, which aligns with the previous research showing how stigmatized vernacular features can be positively reinterpreted by NNSs (e.g., *t*-glottaling by Punjabi immigrants in London studied by Sharma (forthcoming)). In such cases, NNSs may not be aware of the existing stigma, because of a combination of the frequency of a given element in native speech and insufficient opportunities to observe the linguistic behavior of native speakers with regard to that element across different contexts.

NNSs' Sensitivity to Syntactic Placement of rLIKE

We have seen that on the whole, NNSs have acquired the range of NS rLIKE functions and largely reproduce their syntactic distribution. The main goal of the syntactic judgment experiment was to find out whether NNSs can also evaluate the naturalness of rLIKE differently

in different syntactic positions. If they do exhibit gradient judgments, will they be similar to those observed in NSs' responses?

To conduct this comparison, it was necessary to first establish that NSs' judgments would, indeed, depend on the syntactic position of rLIKE. As mentioned above, no published research on US English is available to confirm this. The working assumption was that syntactic position judgments would align with the diachronic hierarchy demonstrated for Toronto English (D'Arcy 2017), especially since Michigan production patterns, as shown in section XX above, coincide with those in Toronto. Some additional support for this assumption comes from a series of in-class experiments on native speaker naturalness judgments of LIKE conducted at the present research site, Michigan State University. Undergraduate students in Suzanne Evans Wagner's general humanities *Language in Society* course have collected hundreds of native US-English-speaker judgments every year since 2011. Participants are asked to judge the naturalness of LIKE in seven different syntactic positions. The results consistently reflect the distributional patterns in the TEA (Wagner, personal communication, 08/22/2017). Nonetheless, I conducted a judgment experiment of my own with NSs at MSU, in order to confirm the working assumption that their judgments would align with the patterns in the TEA. The experiment was also longer and more rigorous than Wagner's class assignment. NNSs were included so that the main research question could be addressed.

Twenty-five NNSs and 31 NSs of English took part in the syntactic judgment experiment (for details on recruitment and the specifics of the presentation of the experiment, see Chapter Three). They listened to 96 sentences and were asked to "decide whether the sentence sounds like something that you might hear in everyday casual speech, for example, on campus". The emphasis on casual speech was important, since it helped reduce the possibility that the

participants could assess the sentences against the prescriptive standards of grammatical correctness. Indeed, judging by the high naturalness ratings given to sentences containing various vernacular elements, including rLIKE, the participants, both NSs and NNSs, understood the instructions correctly.

Nineteen different categories of sentences were included in the experiment, 15 of them featuring rLIKE (see Appendix D). The remaining four categories were the following:

- "Grammatical": authentic (containing some colloquialisms and elements of the vernacular) and grammatically correct sentences taken from IHELP-MI corpus (see details in the Methodology section of this chapter);
- "Ungrammatical": constructed sentences, all grammatically incorrect (the errors reflecting some common errors non-native speakers of English make, such as incorrect word order, lack of subject-verb agreement, pluralized uncountable nouns, etc.);
- "Stigmatized": constructed sentences, containing elements of various stigmatized vernacular varieties of American English, including African American Vernacular English;
- "Unidiomatic": constructed sentences, containing manipulated idiomatic expressions (e.g., "white and black" instead of conventional "black and white").

Validating the Experimental Design. In order to prove the validity of the experiment design, the ratings of the four non-rLIKE categories were analyzed first, in order to make clear predictions with regard to native speaker behavior. Thus, NSs are expected to rate the sentences in the "Grammatical" category highly in naturalness and give low ratings to the "Ungrammatical" and "Unidiomatic" sentences. As for the "Stigmatized", the ratings could vary between the participants depending on their exposure to non-local vernaculars. Generally, given

that the instruction explained naturalness as something that could be heard on campus and that the population of the MSU campus is predominantly white and Michigan-born and raised, and given that prior studies have shown Michiganders to believe their English is 'normal' (Niedzielski, 1999; Niedzielski & Preston, 2000), I would expect the ratings to be substantially lower than those given to the "Grammatical" sentences.

The results supported these predictions. "Grammatical" sentences received high ratings ($M = 4.08$, $SD = 0.43$), "Ungrammatical" sentences received low mean ratings ($M = 1.91$, $SD = 0.45$); small standard deviations for both categories indicate high level of agreement among the participants. "Unidiomatic" sentences low ratings ($M = 2.63$, $SD = 0.88$), that were significantly lower than "Grammatical" sentences ($p < .001$). Somewhat surprisingly, "Stigmatized" sentences were rated even lower in naturalness than the "Ungrammatical" ones ($M = 1.80$, $SD = 0.53$), although a paired-samples *t*-test did not indicate a statistically significant difference ($p = 0.17$).

NNSs also clearly distinguished between "Grammatical" sentences and the other three categories without rLIKE ($p < .001$). Overall, however, the NNSs used a much narrower range of possible responses: they were more generous than NSs in giving 5-point ("perfectly natural" ratings) but avoided giving ratings of 1 or 2 to any sentences. The lowest mean rating given by NNSs to any category was 2.97, while for NSs it was 1.80, an entire point lower. Also, there is an asymmetry in how NSs and NNSs used the scale to assess "Grammatical" sentences and the three other categories, as can be seen in Figure 4.5. For "Grammatical" sentences, the participants of both groups only use the higher end of the scale (3-5 range). For "Ungrammatical" and "Stigmatized", however, NSs almost exclusively used the lower end of the scale (1-3 range), while NNSs only extended the range towards the lower end but did not abandon the higher end completely.

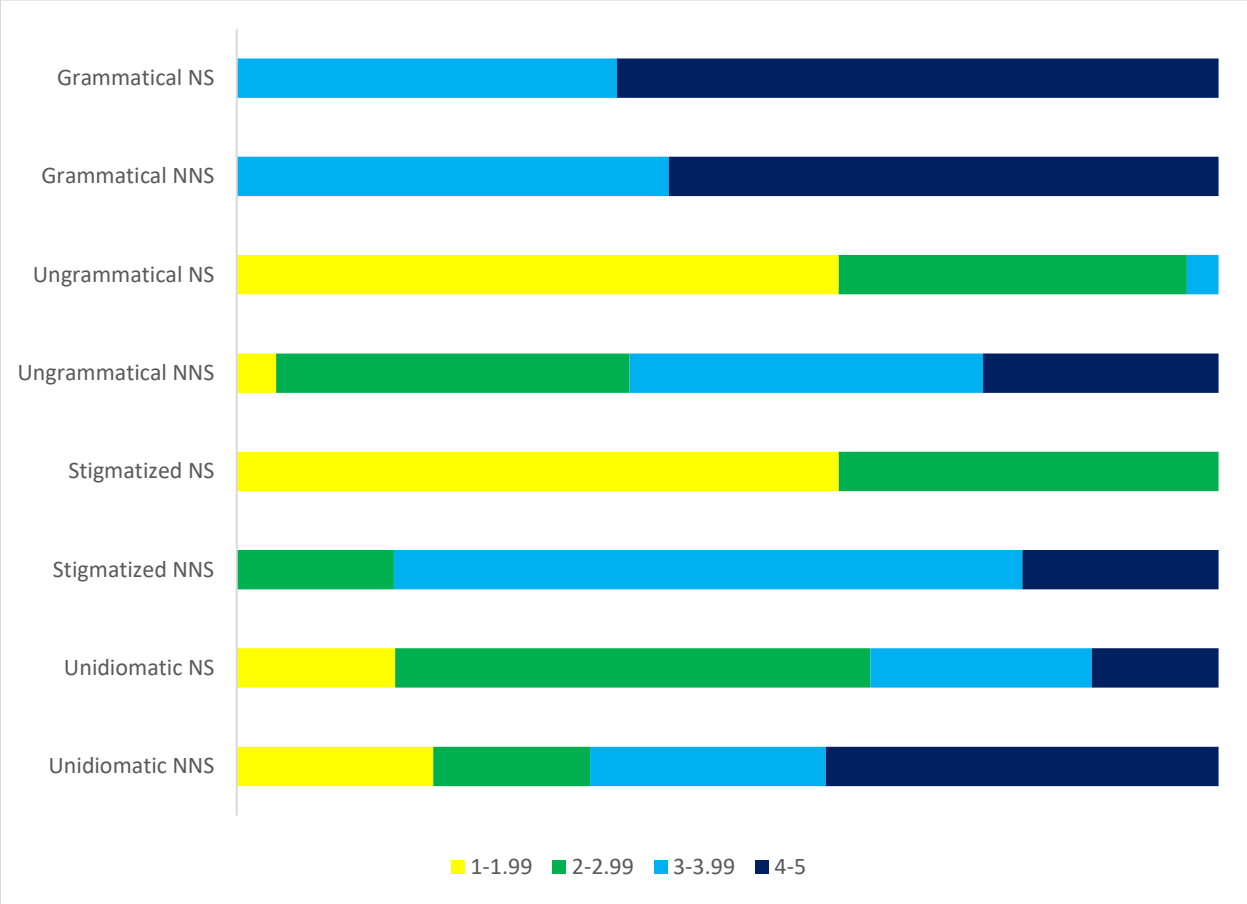

Figure 4.5

Usage of the Scale to Judge "Grammatical", "Ungrammatical", "Stigmatized" and "Unidiomatic" Sentences by NSs (n = x) and NNSs (n = x). The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

The reason for this asymmetry, as well as for the fact that NNS were overall more accepting and tended to give higher ratings, cannot be established with any certainty. However, several speculations can be made. First, it is a well-established fact that non-native grammaticality judgments may differ from native ones, although there are "longstanding concerns about the type of knowledge that judgment data reflect" (see methodological synthesis and meta-analysis in (Plonsky, Marsden, Crowther, Gass, & Spinner (2019); Spinner & Gass (2019)). Second, all sentences were recorded by a female native speaker with a local accent that would have been familiar to the NNSs. She thus sounded native at least phonetically and

intonationally. Therefore, NNSs may have been reluctant to judge the naturalness of her sentences harshly. Finally, the instructions explained naturalness in terms of what can be heard on campus. NNSs are probably exposed frequently on campus to non-native-like constructions used by other NNSs, and thus might have judged such constructions to be natural *for campus*.

It must also be mentioned that not only were the mean ratings of the "Grammatical" sentences given by NSs and NNSs similar, but the variance was rather small, as can be seen in the 95% confidence intervals of the boxplots (Figure 4.6). For the other three categories, the variance was larger, which indicates less uniformity in judgments, i.e., less agreement between participants.

Given this difference in actual scores given together with the similarity of pattern ("Grammatical" being rated around the top of the range and "Ungrammatical" around the bottom), comparing means between groups may not be an appropriate type of analysis, which is the reason why I will only use descriptive statistics and paired-samples *t*-test for comparison of different categories within each group. Most importantly, I will compare the order of naturalness created by each group of participants and find whether the same categories are at the top and at bottom of the respective orders.

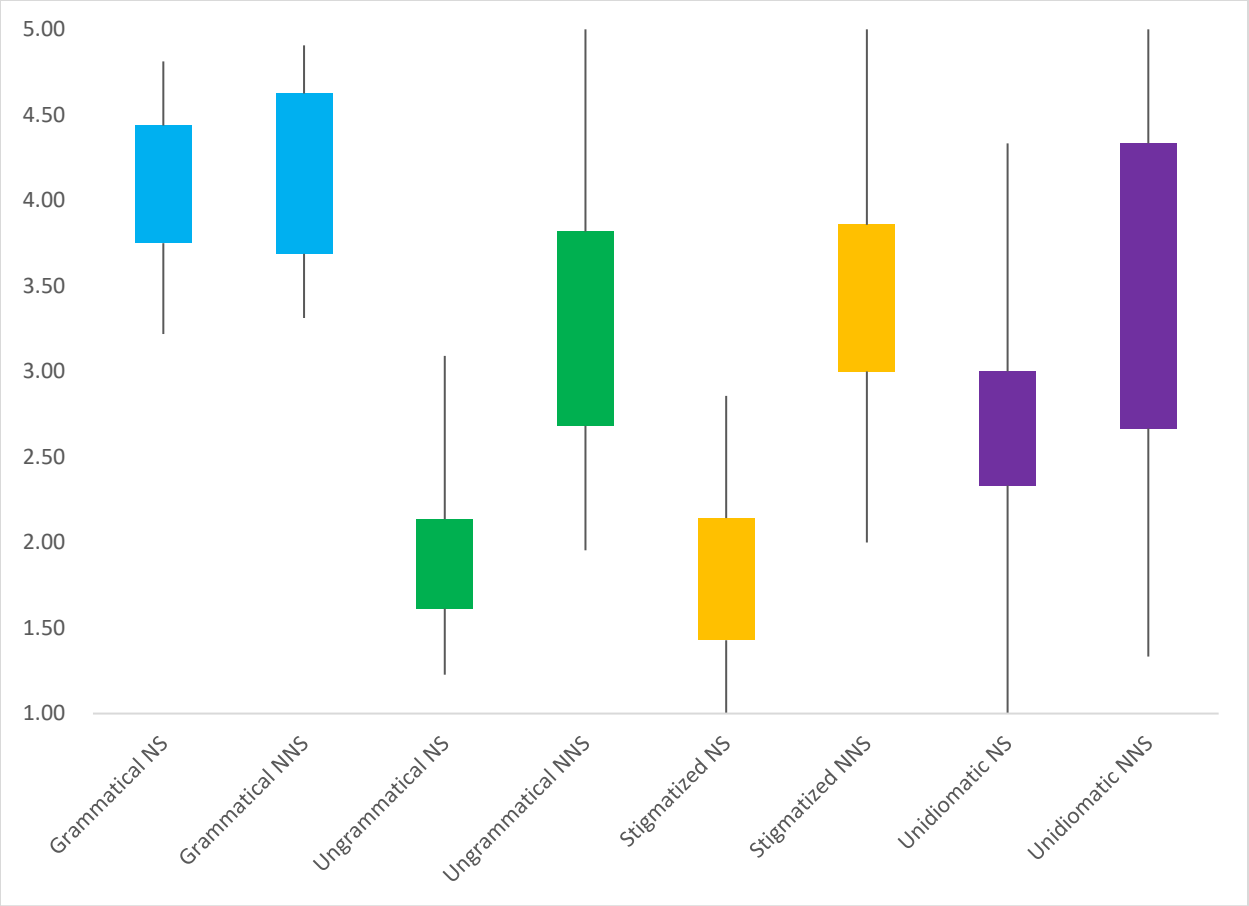

Figure 4.6
Naturalness Judgments of "Grammatical", "Ungrammatical", "Stigmatized", and "Unidiomatic" Sentences by NSs and NNSs. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

Native and Non-Native Judgments of rLIKE. Out of 15 categories of sentences containing rLIKE, 14 represented possible (authentic) usages of rLIKE and one ("Constrained") contained constructed sentences with rLIKE in positions in which, according to D'Arcy (2017) , it does not naturally occur (namely, before a copula or before a personal pronoun). One of the "possible" categories contained an authentic but non-local clause-final rLIKE used mainly in Scotland, Ireland, and some parts of England. Therefore, it is logical to expect that native speakers, most of whom were born and grew up in Michigan, would judge both of these categories as not very natural. If NNSs are sensitive to the existing constraints (both grammatical and regional) on rLIKE usage, they are also expected to give low naturalness ratings. To contrast

truly constrained and possible but rare usages of rLIKE, I included in the experiment the "Rare" category which featured such unusual positions of rLIKE as within a lexeme, before an idiom, or within a long verb phrase; all sentences were taken from D'Arcy (2017). Finally, approximator LIKE (aLIKE) is so well-established in English and so frequently used that NSs are expected to perceive it as highly natural; NNSs with over two years of exposure to authentic speech are expected to do the same.

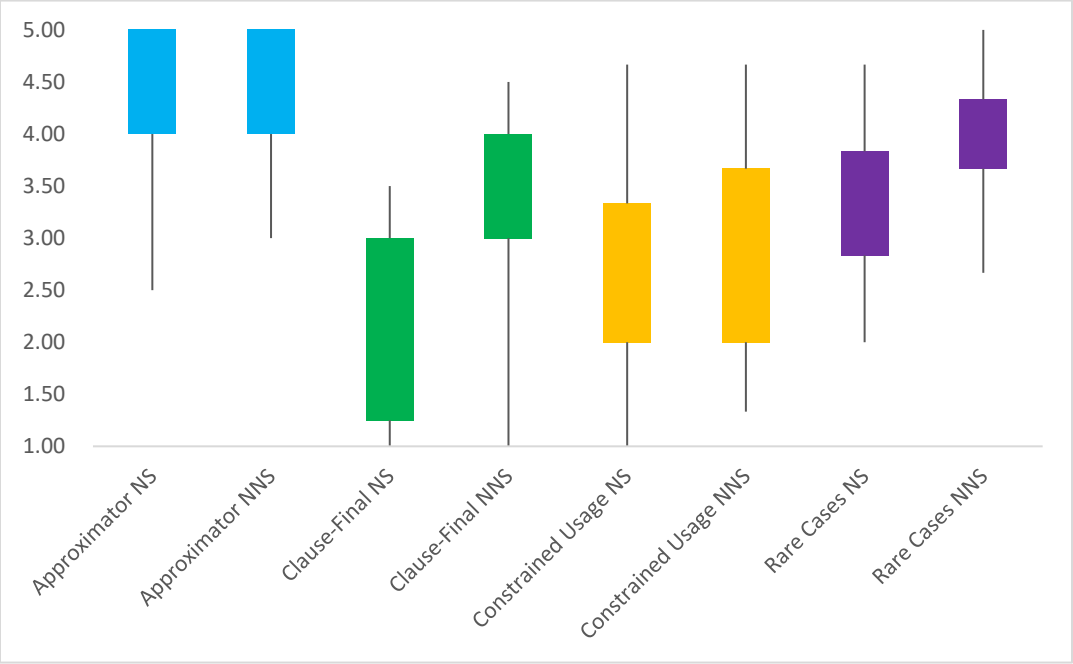

Figure 4.7
Naturalness Judgments of Sentences Containing aLIKE, Clause-final LIKE, Constrained Usages and Rare Cases by NSs and NNSs. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

Indeed, all of the predictions made above were correct. aLIKE was rated as the most natural sounding category by both NSs and NNs. Constrained usages and clause-final LIKE, on the contrary, received low ratings from both groups (see boxplots in Figure 4.7). In fact, the bottom five categories in both groups were the same across groups (although the exact order slightly differed): constrained usages, "Unidiomatic", clause-final LIKE, "Ungrammatical", "Stigmatized" (presented in the order of NS ratings, highest to lowest). At the same time,

authentic rare usages were rated higher than not only these five categories ($p < .001$ for both groups) but some other types of sentences containing rLIKE. Interestingly, the only category for which NSs did not use the entire point range, except the highly ranked aLIKE, was the non-local clause-final LIKE (see Figure 4.8): None of the NSs gave a rating higher than four to this usage, while some gave such rating even to constrained usages.

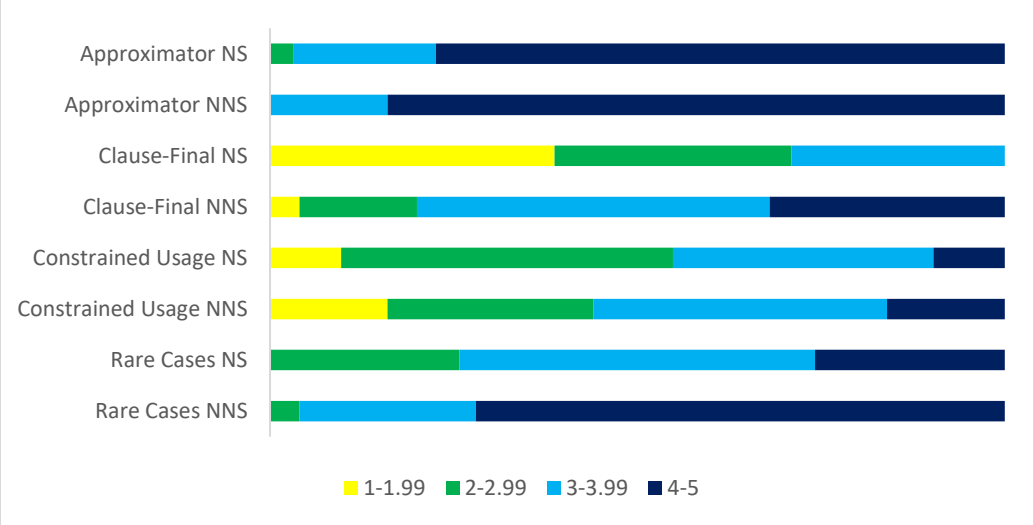

Figure 4.8
Usage of the Scale to Judge aLIKE, Clause-final LIKE, Rare and Constrained Usages of LIKE by NSs and NNSs. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

All of the usages discussed above (with the possible exception of "Stigmatized") can be, with a high degree of certainty, described as either highly frequent or very infrequent (some due to ungrammaticality, some due to various other factors) in the English spoken at MSU campus. Predictably, NSs rated the highly frequent usages (e.g., aLIKE or "Grammatical") as much more natural than infrequent and rare ones (e.g., constrained usages of rLIKE or "Ungrammatical" sentences). NNSs proved to be sensitive to this difference and demonstrated the same response pattern, even though all NNS-given ratings were higher than their NS-given counterparts. Furthermore, NNSs' responses, just as NSs', demonstrated awareness of the difference between truly constrained and rare but authentic usages of rLIKE. However, the most important question,

whether NNSs are sensitive to more subtle differences in frequency of various syntactic positions of rLIKE, still needs to be answered.

Analyzing the order of naturalness created by NNSs and comparing it with the NS-created order may help answer this question. As I have already described, the top category and the bottom five categories are the same in both groups. In the middle, however, the differences are notable (see Table 4.4).

Table 4.4
Order of Naturalness of Stimuli in Syntactic Judgment Experiment Established by NSs and NNSs

Rank	NS			NNS		
	Function of rLIKE	Mean	SD	Function of rLIKE	Mean	SD
1	Approximator LIKE	4.26	0.70	Approximator LIKE	4.44	0.62
2	Grammatical	4.08	0.71	pLIKE within to-infinitive	4.34	0.79
3	pLIKE within to-infinitive	3.77	0.85	pLIKE before to-infinitive	4.20	0.84
4	pLIKE to the left of main verb	3.65	0.98	Grammatical	4.17	0.53
5	mLIKE at subordinate CP	3.54	0.72	mLIKE at matrix CP / non sentence-initial	4.16	0.81
6	pLIKE: DP as complement of PP	3.53	0.98	pLIKE: DP as argument of vP	4.12	1.00
7	mLIKE at matrix CP / non sentence-initial	3.52	0.78	Rare usages of rLIKE	4.07	0.55
8	mLIKE at matrix CP / sentence-initial	3.47	0.96	mLIKE at subordinate CP	3.97	0.76
9	pLIKE: DP as argument of vP	3.34	1.15	pLIKE between modal and main verbs	3.96	0.79
10	Rare usages of rLIKE	3.32	0.66	pLIKE to the left of main verb	3.92	0.89
11	mLIKE at subordinate TP	3.11	0.98	pLIKE: DP as complement of PP	3.90	0.88
12	pLIKE: nP	3.05	0.85	mLIKE at matrix CP / sentence-initial	3.88	1.11
13	pLIKE between modal and main verbs	3.07	0.99	mLIKE at subordinate TP	3.70	0.80
14	pLIKE before to-infinitive	2.80	0.95	pLIKE: nP	3.62	0.99

Table 4.4 (cont'd)

15	Constrained usages of rLIKE	2.73	0.85	Stigmatized	3.52	0.82
16	Unidiomatic	2.63	0.88	Ungrammatical	3.38	0.92
17	Clause-final LIKE	2.06	0.86	Unidiomatic	3.30	1.17
18	Ungrammatical	1.91	0.45	Clause-final LIKE	3.26	0.88
19	Stigmatized	1.80	0.53	Constrained usages of rLIKE	2.97	0.97

Note. Gray area indicates the bottom five categories.

In Figure 4.9, all rLIKE categories are plotted on a line graph, with two lines each of which represents one participant group, NSs and NNSs. The order from right to left is generally the native order of naturalness, although I combined contexts in two instances (sentence-initial and non-sentence-initial matrix CP and DP context as argument of vP and complement of PP). As can be seen in the graph, while almost all NNS-given naturalness ratings are higher, the response patterns are generally similar, with aLIKE being rated very high in naturalness and non-local clause-final LIKE and constrained usages rated lower than other types of sentences. However, it is necessary to analyze the judgments of rLIKE in various contexts more closely, which can be done by dividing the categories into three major domains: clausal, nominal, and verbal.

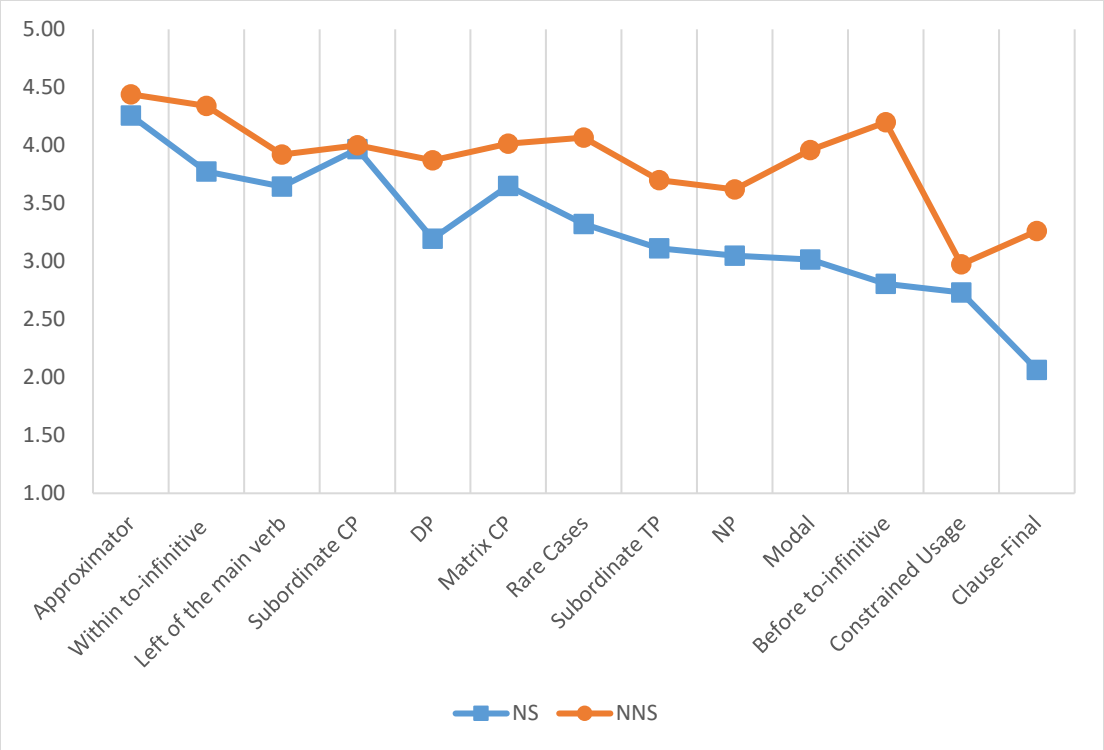

Figure 4.9

Naturalness Judgments of All rLIKE Categories by Native (n = 31) and Non-Native (n = 26) Speakers of English. The DP category includes two functional contexts of a DP, argument of vP and complement of PP; the matrix CP category includes both sentence-initial and non-sentence initial mLIKE in the CP matrix CP contexts. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

*Clausal Domain***.** From the diachronic perspective, the appearance of rLIKE in matrix CP contexts began earlier, and, therefore, is more established than rLIKE in subordinate CP. As the distribution of mLIKE was "identical to the result for matrix CP" (D'Arcy, 2017, p. 122) in the subset of Toronto English Corpus that D'Arcy used in her study, no strong hierarchy between these contexts should be expected in the NS judgment data. Interestingly, however, NSs judged rLIKE in subordinate CP context as more natural than in matrix CP ($p = .01$), regardless of whether rLIKE appeared in sentence-initial or non sentence-initial matrix CP context (the difference between those categories is not significant and they take adjacent position in the table). According to D'Arcy, the most recent position available to mLIKE is at subordinate TP (i.e., following the subordinator and not preceding it), and NSs did rate the naturalness of

sentences with mLIKE in this position much lower than that of mLIKE at subordinate CP or matrix CP (p <.001); see boxplots in Figure 4.10.

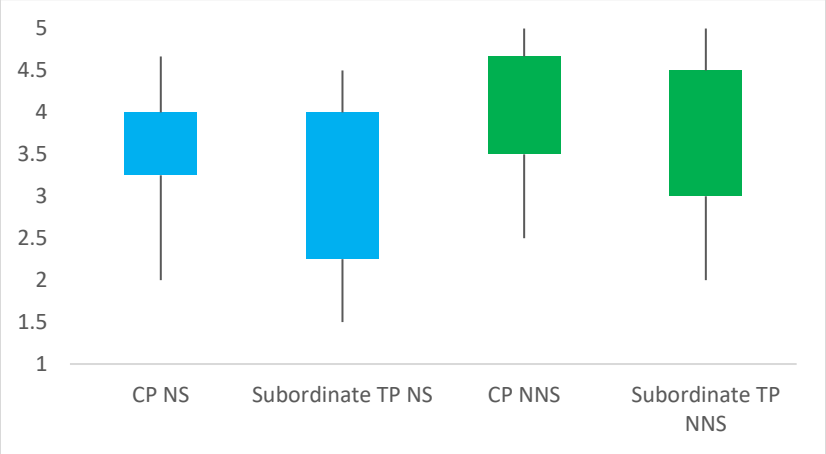

Figure 4.10
Naturalness Ratings of mLIKE in Subordinate CP and Subordinate TP Contexts by Native and Non-Native Speakers of English. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

NNSs rated rLIKE in the subordinate TP context lower than in the subordinate CP context, although the difference was not as prominent and did not reach the threshold of statistical significance (p = 0.07); see boxplots in Figure 4.10. Furthermore, the patterns of individual variation in rating these two contexts were similar among NSs and NNSs (see Figure 4.11), with most, but notably not all listeners rating subordinate CP contexts as more natural.

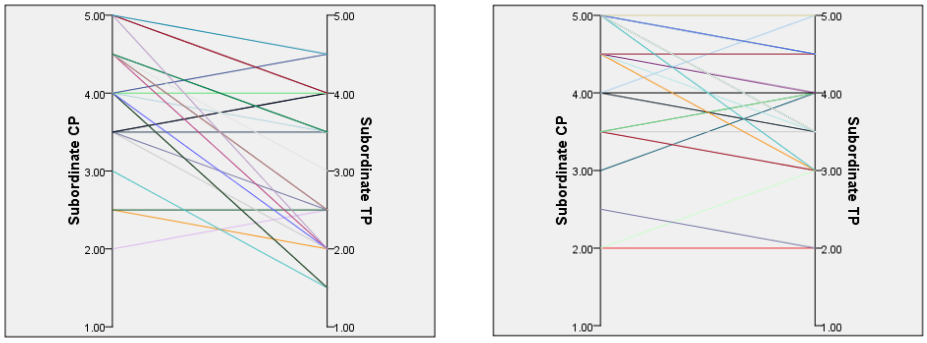

Figure 4.11
Individual Variation in Naturalness Ratings of rLIKE in Subordinate CP and Subordinate TP Contexts by Native and Non-Native Speakers of English. NS ratings are on the left. Each line represents one participant. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

While the hierarchy of mLIKE (clausal) contexts created by NS and NNS judgments was not exactly the same as the one developed by NS judgments, with sentence-initial and non sentence-initial CP contexts being rated differently (non sentence-initial CP as the most natural, then subordinate CP, and then sentence-initial CP right above the subordinate TP category), this order (if we ignore the difference in ratings given to mLIKE in sentence-initial and non-sentence-initial matrix CP contexts) generally matches the order of frequency with which NSs and NNSs use marker LIKE (matrix CP, subordinate CP, subordinate TP).

Nominal Domain. Within the nominal domain, different ratings of pLIKE in the DP and nP contexts can be expected, as nP context is the newest available for pLIKE in this domain. Indeed, the nP naturalness ratings are significantly lower ($p < .001$); see boxplots in Figure 4.12. This matches the production data for both groups of speakers who use pLIKE in the DP context with much higher frequency than in the nP context.

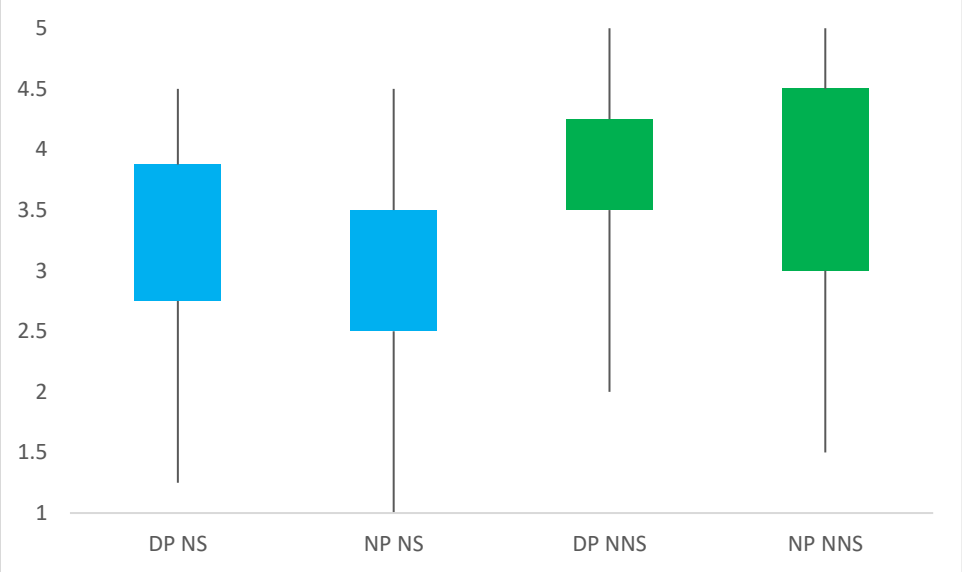

Figure 4.12
Naturalness Ratings of rLIKE in DP and nP Contexts by Native and Non-Native Speakers of English. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

However, if two positions of DP (complement of PP and argument of vP) are analyzed separately, only DP as the complement of PP would be rated significantly higher than nP. For NNSs, the opposite is true: only DP as the argument of vP were rated significantly higher than nP (see boxplots in Figure 4.13). Production data were not coded with regard to these positions, so it is impossible to compare the distributions with judgment data.

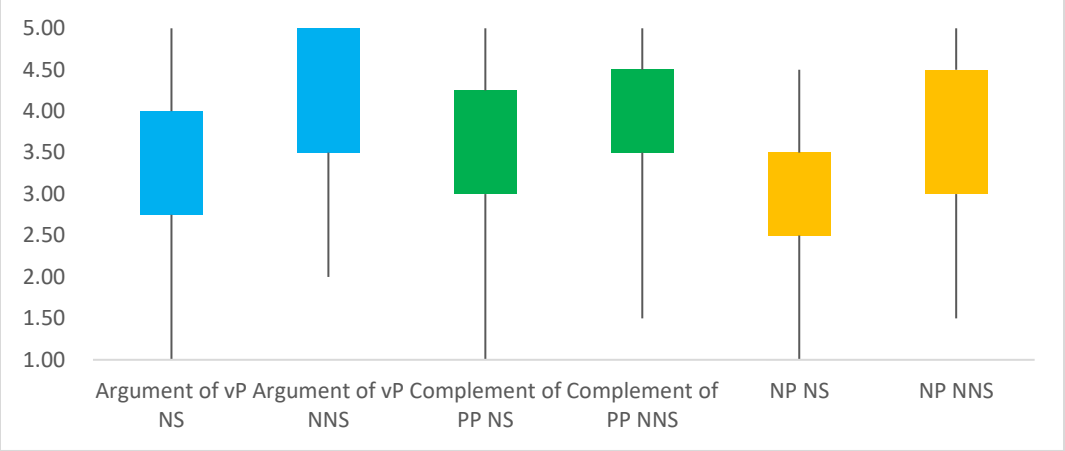

Figure 4.13
Naturalness Ratings of rLIKE in Different DP (argument of vP and complement of PP) and nP Contexts by Native and Non-Native Speakers of English. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

Verbal Domain. Within the verbal domain, as expected, pLIKE on the left periphery of the main verb, the oldest context available for pLIKE in a sentence, was rated high in naturalness by native speakers of English. NSs rated pLIKE within *to*-infinitive even higher (but not significantly so) than pLIKE on the left periphery of the main verb, while pLIKE before *to*-infinitive received the lowest naturalness rating among all authentic and local rLIKE categories, which is consistent with one of the constraints on pLIKE usage established by D'Arcy (namely, that pLIKE occurs on the immediate left periphery of the lexical verb if other verbs (e.g., auxiliary or modal) or *to* were present). pLIKE between a modal verb and a main verb also received low mean naturalness rating from NSs, although the variance was high (see boxplots in Figure 4.14).

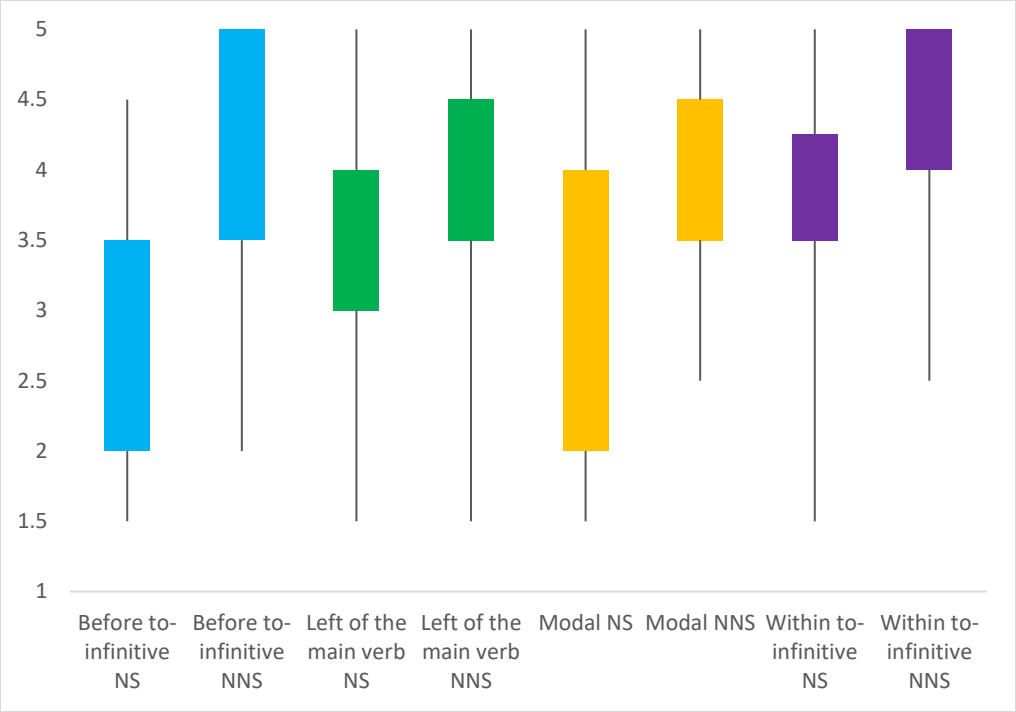

Figure 4.14
Naturalness Ratings of pLIKE in the Verbal Domain Contexts by Native and Non-Native Speakers of English. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

NNSs, on the other hand, rated pLIKE both before and within *to*-infinitive highly in naturalness, and similarly so (these categories occupy adjacent positions in the hierarchy, see Figure 4.14). Also, pLIKE on the left periphery of the main verb received relatively low ratings, even lower than pLIKE between a modal and a main verb. This suggests that the native patterns of pLIKE usage in the verbal domain may either be more difficult to acquire than the patterns of the clausal and nominal domains or require longer exposure to native English speech.

Only five NNSs (two males, three females) exhibited native-like or almost native-like response patterns with regard to the verbal domain (see Figure 4.15). The criteria upon which I considered a pattern to be native-like were simple: high naturalness ratings of pLIKe within *to*-infinitive and on the left periphery of the single main verb and lower (compared to the previous two categories) ratings of pLIKE between a modal and a main verb and before *to*-infinitive.

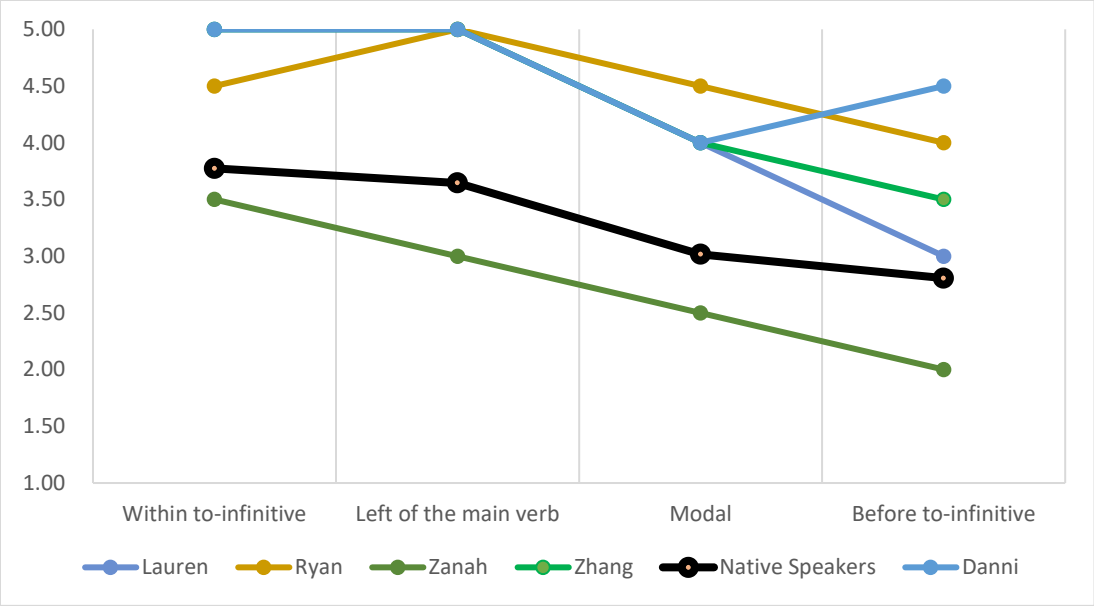

Figure 4.15

Native-like Response Patterns within the Verbal Domain. The rating of 1 means "not natural at all", the rating of 5 means "perfectly natural".

Four of the five NNSs followed the native pattern exactly, giving almost the same rating to both highly natural categories or judging pLIKE on the left of the main verb slightly lower than within *to*-infinitive and considering pLIKE between a modal and a main verb as more natural than before *to*-infinitive. One NNS (Danni) did not demonstrate these fine-grained differences but, crucially, judged the first two categories much higher in naturalness than the other two.

These five NNSs came from four different countries and pursue five different majors. However, four of them (Lauren, Ryan, Zhang, and Danni) reported longer than average (mean = 40.08 months) length of residence in the United States: 53, 84, 53, and 41 months, respectively. Only Zanah resided in the United States for a shorter period (29 months) prior to taking part in the experiment. Three of them (Zanah, Zhang, and Danni) are heavy pLIKE users themselves, with frequencies of 20, 27, and 35 (respectively) tokens of pLIKE per 1000 word, which is well above the group average of 14. Neither of these characteristics (longer than average LOR, high

frequency of pLIKE usage), however, describe these NNSs exclusively: four other NNSs reported LOR much longer than average and another five reported the same LOR as Danni (41 months); five other NNSs demonstrated above average frequency of pLIKE usage. Therefore, while length of residence may contribute to the acquisition of native-like patterns of judgment with regard to pLIKE in the verbal domain, it is unlikely to be the only contributing factor, and it certainly does not guarantee such acquisition. Heavy pLIKE usage does not necessarily coincide with native-like judgment patterns as well.

Conclusion. NNSs of English in this experiment demonstrated remarkable sensitivity towards the native patterns of rLIKE usage by rating the most established and frequent usages (e.g., aLIKE) higher in naturalness and constrained and newer and/or rare usages of rLIKE (e.g., pLIKE in subordinate TP or nP contexts). These results are further supported by the consistency of NS and NNS ratings of other types of sentences without rLIKE: While authentic and grammatically correct sentences were rated highly by both NS and NNS participants, constructed sentences containing grammatical errors, as well as sentences with the elements of stigmatized vernacular or inverted (manipulated) idioms, received unanimously low ratings. However, the similarity between native and non-native judgments of rLIKE seems to be restricted to aLIKE, mLIKE and the nominal domain of pLIKE. Within the verbal domain, the NNSs's responses do not align with the native judgments; crucially, one of the most productive adjunction sites for pLIKE, the left periphery of the main verb, is not recognized as such by NNSs, while they do not distinguish between the placement of pLIKE within or before *to*-infinitive (the latter being an almost constrained context, the former highly productive and rated as natural by NSs). While several NNSs demonstrated native-like pattern of pLIKE judgment, no other distinct characteristic or characteristics distinguish them from the other participants. Thus, the current

data do not allow to determine the reasons why the acquisition of a native-like pattern of pLIKE judgments is less successful within the verbal domain, although it is possible that length of residence may be one of the contributing factors, as four of the five participants who exhibited such pattern also reported longer than average LOR.

Conclusion

Several major findings resulted from my study of non-native usage of rLIKE and patterns of syntactic judgment. First, all of the NNSs in the sample, including those who spent less than two years in the United States and thus in the target language environment, produced at least a few tokens of LIKE, thus showing that it was present in their discourse repertoire. As a group, NNSs demonstrated striking nativelikeness in the way they used different functions of LIKE and in syntactic placement of clause-medial LIKE in their own speech. Importantly, as the reference group of native speakers used in this study consisted of students from the same university who also grew up in the area, it can be inferred that the cross-function distributions found in NNSs mirrored the patterns found in the speech these NNSs had been exposed to prior to data collection, i.e., in the input these NNSs received. Given that previous research found differences in preferred patterns of LIKE placement between different varieties of English (Schweinberger, 2011), this level of similarity between native and non-native patterns is unlikely to be coincidental.

Native patterns of syntactic judgments, that is, the hierarchy of naturalness of rLIKE functions constructed on the basis of native speaker responses, also appeared to have been generally acquired by NNSs. Closer analysis revealed that, while longer length of residence, a factor often found to be important in the LIKE research (Diskin, 2013, 2017; Hellermann & Vergun, 2007; Liu, 2016), was not directly contributing to higher rates of rLIKE usage, it may be

affecting syntactic judgments. Judgment of the placement of LIKE in the verbal domain was the only area in which NNSs as a group differed notably differed from their native-speaking peers, but five NNSs appeared to have acquired the native pattern of judgment in this domain. All of these NNSs resided in the U.S. for over two years, and LOR of four of them was over the rather high average benchmark of 40 months. Even though above-average LOR did not guarantee the native-likeness of judgment patterns, it may serve as indirect evidence that, in combination with other factors, longer LOR (which, in case of undergraduate students, also implies daily exposure to native speech) may facilitate acquisition of native patterns.

Table 4.5

rLIKE Usage (Token Number and Normalized Rates) across Speakers and rLIKE Functions

Name	Gender	Country Group	Friends	Intentions	Americanness	rLIKE (per 1,000 w.)	mLIKE (per 1,000 w.)	aLIKE (per 1,000 w.)	pLIKE (per 1,000 w.)	qLIKE (per 1,000 w.)
Zhang	m	2	0	3	yes	59 (78.25)	24 (31.83)	2 (2.65)	20 (26.53)	4 (5.31)
Juju	f	3	1	0	yes	84 (76.57)	43 (39.2)	1 (0.91)	36 (32.82)	0 (0)
Danni	f	1	-	2	yes	103 (74.85)	33 (23.98)	2 (1.45)	48 (34.88)	15 (10.9)
Aisya	f	1	2	2	yes	49 (50)	19 (19.39)	0 (0)	23 (23.47)	7 (7.14)
Iris	f	2	-	3	yes	51 (47.53)	25 (23.3)	2 (1.86)	18 (16.78)	1 (0.93)
Emily	f	1	2	2	yes	51 (40.64)	8 (6.37)	7 (5.58)	30 (23.9)	4 (3.19)
Zanah	f	1	0	3	(yes)	41 (34.63)	5 (4.22)	0 (0)	24 (20.27)	1 (0.84)
Ryan	m	1	2	3	no	53 (32.92)	17 (10.56)	1 (0.62)	13 (8.07)	15 (9.32)
Maxwell	m	1	1	0	no	22 (30.39)	12 (16.57)	2 (2.76)	5 (6.91)	0 (0)
Anna	f	2	1	1	yes	27 (30.2)	5 (5.59)	4 (4.47)	16 (17.9)	0 (0)
Su	f	3	2	0	yes	50 (27.44)	17 (9.33)	1 (0.55)	20 (10.98)	8 (4.39)
CY	m	2	2	2	no	17 (19.36)	3 (3.42)	2 (2.28)	8 (9.11)	1 (1.14)
Katy	f	1	2	0	no	17 (18.68)	2 (2.2)	2 (2.2)	13 (14.29)	1 (1.1)
Maggie	f	2	0	1	yes	11 (18.68)	4 (6.79)	2 (3.4)	5 (8.49)	0 (0)
Marsha	f	1	2	-	(yes)	16 (16.97)	7 (7.42)	0 (0)	5 (5.3)	1 (1.06)
Ibra	m	3	2	0	no	27 (16.81)	3 (1.87)	2 (1.25)	16 (9.96)	4 (2.49)
Carlos	m	4	2	-	yes	28 (16.4)	11 (6.44)	2 (1.17)	7 (4.1)	8 (4.69)
Frank	m	4	2	2	(yes)	14 (12.28)	2 (1.75)	1 (0.88)	11 (9.65)	0 (0)
Lauren	f	1	2	3	(yes)	12 (11.39)	4 (3.8)	1 (0.95)	7 (6.64)	0 (0)
Griggs	m	2	3	2	(yes)	15 (9.55)	3 (1.91)	1 (0.64)	10 (6.37)	0 (0)
Ange	f	3	1	1	yes	12 (9.35)	6 (4.67)	0 (0)	2 (1.56)	4 (3.12)
MT	m	2	0	0	no	6 (7.64)	1 (1.27)	3 (3.82)	2 (2.55)	0 (0)
Jenny	f	2	2	3	no	5 (5.69)	1 (1.14)	2 (2.28)	2 (2.28)	0 (0)
An	f	2	3	2	(yes)	2 (2.7)	1 (1.35)	0 (0)	1 (1.35)	0 (0)
John	m	4	2	3	no	2 (2.22)	0 (0)	0 (0)	2 (2.22)	0 (0)

Note. The table is sorted following the normalized rates of rLIKE usage (per 1,000 words). Highlighted rows represent top and bottom rLIKE users. The following codes applied. For Country Group: 1 = Southeast Asia, 2 = China and Taiwan, 3 = Africa, 4 = Other. For Friendship: 0 = person socializes mostly with people from their home country; 1 = person socializes mostly with other international students; 2 = person has a diverse circle of friends; 3 = person socialized mostly with Americans. For Intention: 0 =

Table 4.5 (cont'd)
person intends to return to their home country upon graduation; 1 = person would prefer to stay and work in the U.S. for a while but later return to their home country; 2 = person would like to live in a third country; 3 = the person would like to stay in the U.S. For Americanness: yes = person recognizes LIKE as a specifically American phenomenon; (yes) = person recognizes LIKE as a phenomenon characteristic of a particular group of Americans; no = person does not associate LIKE with Americanness specifically.

The summary of rLIKE usage across individual speakers and functions of rLIKE is presented in Table 4.5. Despite the remarkable native-likeness of their group behavior with regard to rLIKE usage and judgments of its syntactic placement, a high degree of variation between individual speakers is evident from the table. It can also be seen how diverse this sample is in terms of speaker gender, origin, friendship circles, post-graduation plans, and attitudes towards LIKE. Gender, a widely-studied factor in rLIKE research (e.g., Nestor et al., 2012; Truesdale & Meyerhoff, 2015), appeared to influence rLIKE usage by the participants of this study as well: overall, females used more rLIKE than men. By-function analysis, however, revealed that the overall female lead stems mainly from their high rates of using clause-medial particle LIKE (women in this sample used it almost twice as frequently as their male peers). The female lead in the usage of clause-initial marker is only slight, while no clear differences between men and women are seen with regard to their usage of approximator and quotative LIKE. This reflects the overall gender distributions in American English reported in the literature.

One of the most important findings of this study was the connection between perceiving LIKE as a symbol, or a signal, of Americanness. As shown by various researchers, NNSs may unconsciously overuse socially salient vernacular features as a way to signal nativeness (Britishness, Frenchness, etc., depending on the context of the study) in their own speech (Cheshire et al., 2011; Regan et al., 2009; Sharma, forthcoming, 2016). It seems to be the case with the heaviest rLIKE users in my sample: Regardless of their overtly proclaimed attitudes towards LIKE, which may be even negative, they all recognize rLIKE as a specifically American phenomenon. The non-users and light users, on the other hand, are much more likely to not associate LIKE with Americanness at all, or to associate it with a particular group of Americans.

CHAPTER FIVE

"IT'S AN AMERICAN THING": NATIVE AND NON-NATIVE BELIEFS ABOUT AND ATTITUDES TOWARDS REMARKABLE LIKE

In this chapter, I will explore the beliefs, thoughts, and opinions about rLIKE that native- and non-native-speaking participants of this study shared with me, as well as the attitudes towards rLIKE they inadvertently exhibited during the matched-guise experiment. However, before presenting the findings of my study, I will describe the tasks and materials/instruments I used to collect the data, as well as the processes of data treatment and analysis. The context of the study and the overall procedure were described in Chapter Three; the participants were introduced in that chapter as well.

Methodology

Matched-Guise Experiment

The main goal of the matched-guise experiment was to elicit beliefs about and attitudes towards the various remarkable functions of LIKE in authentic speech. The technique itself was developed by Lambert, Hodgson, Gardner, and Fillenbaum (1960) and has been widely used in sociolinguistic research since then. The essence of the technique is that two versions of the same speech sample are created in a way that only one characteristic (e.g., the frontedness of a particular vowel or the presence of a particular discourse marker) distinguishes the version; this can be accomplished by digitally manipulating the same recording or recording the same speaker twice (controlling for intonation, speech rate and other factors to the best of the speaker's ability).

The stories used as a basis for the matched-guise task were taken from the IHELP-MI corpus (Wagner, Mason, Nesbitt, Pevan, & Savage, 2016). Six stories, each from a different

speaker (all white, 18-22 years old, two males, four females), were selected. The term "story" is used loosely and refers to a coherent and semantically bounded segment of the interview in which the speaker discussed a specific topic or event. It is true that the topic may affect perceptions; for example, Campbell-Kibler (2007), in a matched-guise study of the perceptions of (ING), found that "the content of the recording strongly impacted perceptions of the speaker" (p. 34), with one of the speakers judged as "lazy" almost 90% of the time, since in the story he complains how much effort it takes to attend a movie. However, as Campbell-Kibler (2009) noted in a later article, "there is no such thing as truly neutral content and in seeking it, we are likely not only to fail but to sacrifice important insights about the complex interplay between content and form" (p. 138). Thus, I decided to use a range of topics, from relatively neutral descriptive topics to the discussion of career and/or profession, to a very personal account of family bonds. Below are two examples of such stories. The first one, told by a woman, was not an actual story in the sense of a personal narrative; the speaker was talking about the difference in meaning between two words, "cap" and "hat". The second story, told by a man, was an account of his career path from a "lights guy" to a stage manager in a school or college theater. The number assigned to the stories reflects the order in which the stories were first presented to the listeners. All six stories can be found in Appendix E.

Story 1. A cap and a hat

Um, I feel like a cap is a -- more specific. A cap is a type of hat. Like a hat could be all these different sorts of things that you put on your head. Like it wouldn't be like a blanket that you put on your head but like different things that you can put on your head to cover your head. And a cap would be, I would think more like a baseball cap. Or, um, like one

of those grandpa's caps. Yeah. Because a hat is much broader. So you could use different things, and caps, usually like baseball caps are probably made out of the same thing.

Story 5. Stage lighting

Well, my brother is an actor. And he likes to do the theatre. So um, he told me that Mrs. James[10], the director needed someone to do lights my freshman year. And I was like, "Okay, I'll go see what it's about." It was totally random. And I just thought -- I liked doing lights. I did two shows. 'Cause there's a musical, and a play every year. So I did lights for both of them and then, um, the next year she just came up to me. She's like, "We need a stage manager. The other one left." So I was like, "Okay." I di- +didn't -- so I had no idea what I was doing. Like I didn't even know what it was I was supposed to do. because I was a part of the last productions but I – like I was just the light guy. I didn't know very many people uh yeah. And, like 'cause you're supposed to, uh, at first I didn't know like I'm supposed to direct other people to do stuff. So I was trying to do everything myself. And it took a long time to learn. But by my senior year I was -- I was teaching bunches of people how to do it.

Two guises of each story were prepared. The LIKE guise was the original version of the story as told by the speaker. The bare guise was the same story with all instances of rLIKE removed (qLIKE replaced). The unremarkable LIKEs were not removed, as Maddeaux and Dinkin (2017) demonstrated their lack of salience. Both versions of each story were recorded by volunteer actors (the gender of the actor matched with the gender of the speaker) who listened to the original recordings and were asked to mimic the intonations of the original speakers. The decision to use actors instead of digital manipulation of the recordings was made for two reasons.

[10] In case a name was mentioned in a story, it was replaced with a pseudonym matched to the original in the number of syllables and word stress. “Mrs. James” is an example of such pseudonym.

First, the original recordings were of poor quality and contained a lot of background noise. Second, it was important to prevent the participants from perceiving bare guises as technologically manipulated/inauthentic because of the length of pauses and disrupted intonation contours.

In the first part of the experiment (*matched-guise profiling task*), the participants (NSs and NNSs) were asked to guess the age of and freely profile each speaker. Garrett, Williams, and Evans (2005) used a similar technique ("keywords technique", in their terms) to investigate the attitudes of native speakers of each of the different varieties of English (e.g., US English, British English, etc.) to the other varieties. They argued that the task is "meaningful to the respondents" and "reflects the most salient ways in which they make their judgments" (p. 216). The exact instruction given to the participants was the following: "Please try to guess the speaker's age and write everything you can say about the person who told the story. Phrases, words divided by a comma, or complete sentences are all fine: choose the style that is most comfortable for you." For NSs, the task was incorporated into the *Qualtrics* survey software so that they just moved on to it after completing the background survey. To NNSs the task was presented via PsychoPy software (Peirce, 2007) on my laptop. NNSs were given a headset and a mouse to move through the task but were asked to respond on paper (a set of numbered sheets and pencils were provided).

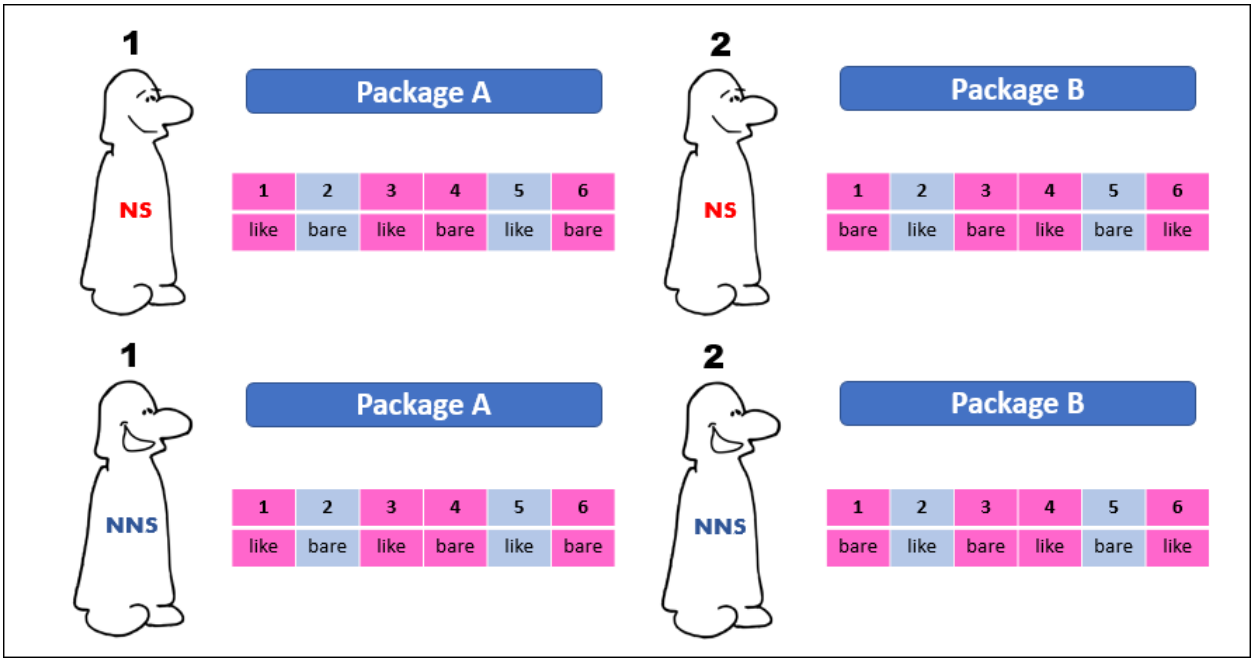

Figure 5.1
Distribution of Stories in the Packages for the Matched-guise Experiment. Pink cells represent stories narrated by a female voice, blue cells – by a male voice.

All participants heard all six stories, but each in one guise only. Two packages were formed to be then randomly presented to a participant (see Figure 5.1). Package A included the LIKE guises of the stories "A Cap and a Hat" (1), "Parting with the Family" (3), "Stage Lighting" (5), and bare guises of the stories "A Townhouse and a Duplex" (2), "Future Career" (4), "A Haircut" (6). Package B included the LIKE guises of "A Townhouse and a Duplex", "Future Career", "A Haircut", and bare guises of "A Cap and a Hat", "Parting with the Family", "Stage Lighting".

In the second part of the experiment (*matched-guise attitudinal task*), participants heard the same stories in the same guises again, although this time in randomized order (via *PsychoPy* for NNSs and via *Qualtrics* for NSs). However, the task was different. The participants were given a list of seven adjectives and the following instruction: "Please decide how well each of the adjectives listed below describes the speaker." A five-point Likert scale with labels was used; the labels were the following: "Describes the speaker perfectly"; "Describes the speaker

somewhat well"; "I am not sure"; "Doesn't describe the speaker well"; "Doesn't describe the speaker at all". The adjectives for the task were selected as the most common ones used in previous attitudinal studies of rLIKE (Buchstaller, 2006; Dailey-O'Cain, 2000; Maddeaux & Dinkin, 2017): *friendly, intelligent, educated, polite, ambitious, attractive, confident*. To make sure that NNSs would not experience any lexical difficulty, the adjectives were tested on a group of NNS students at the English Language Center who were asked to freely define each of them. As those students were not yet allowed to be enrolled in mainstream classes at MSU due to insufficient English proficiency, it would be safe to assume that if these adjectives were not problematic for them, they would not be problematic for students with higher English proficiency as well. Indeed, none of these adjectives proved to be lexically challenging neither to ESL students nor to the actual participants of the study.

Attitudinal Interview (NNSs)

First, the participants were given two of the written profiles they created during the matched-guise profiling task and asked to listen to the recordings again and comment upon those profiles. For example, if a participant described a speaker as a "party girl", the participant was asked to elaborate, explain what they meant by that and what exactly made them think about that speaker in that particular way. I selected the profiles for discussion while the participants were working on the *syntactic judgment task*.

After that, the participants were asked if they were able to figure out the goal of the study. Then they were explicitly told that the study was about rLIKE (I used the phrases "conversational LIKE", "not as a verb", and gave examples), and asked a series of questions about it. The questions were concerned with the saliency of rLIKE for NNSs, rLIKE meanings,

patterns of usage and factors upon which the decision to use or avoid rLIKE can be based, etc. (see Appendix F for the interview protocol).

Attitudinal Survey (NSs)

The survey, administered via *Qualtrics*, opened with a question about the purpose of the study, proceeded to a simple explanation and examples, and ended with a series of mostly open-ended questions about rLIKE, concerning its saliency, patterns and frequency of usage, etc. At the end of the survey the participants were asked to add any additional information or opinions if they wanted to (see Appendix G for the survey).

Data Coding and Analysis

Attitudinal Data. Beliefs about rLIKE were analyzed qualitatively. Transcripts of attitudinal interviews with NNSs were coded for themes and then, after several prominent themes emerged, coded anew with regard to those themes. For example, after I noticed that the theme "LIKE is an American thing" began to emerge in the data, I reanalyzed all the interviews looking specifically for any words or phrases related to that theme.

As NSs were not interviewed but responded to an online survey, themes were largely predetermined by the questions. However, close reading revealed that participants tended to return to certain themes across various questions, so additional layer of thematic coding was added. The procedure was similar to that used with the interviews: After a theme would emerge, I would re-analyze all answers with regard to it.

Experimental Data. Finally, the results of the matched-guise experiment were analyzed. With the *profiling task* data, I followed the content analytic procedure developed by Garrett et al. (2005). First, a number of descriptive themes were identified. For example, if a participant wrote that a certain speaker "sounds smart", it was coded as [+intelligence], and if the statement was

the opposite (e.g., the speaker "doesn't sound smart at all"), the code [-intelligence] was used. Only explicit statements were coded, that is, if a participant did not provide any comment with regard to the intelligence of the speaker, no code for intelligence was applied. After that, the descriptors were grouped into larger themes (i.e., "personal characteristics", "physical appearance", etc.). This procedure worked for the NS data but could not be fully applied to NNS responses due to lack of quality data (see more later in this chapter). Finally, six composite profiles were created using the descriptors provided by the participants: two across-guise profiles (by NSs and NNSs), two LIKE guise profiles, and two bare guise profiles.

The results of the second part of the experiment were analyzed by a Repeated Measures ANOVA with rating scores as a dependent variable, Group and Guise as between-subject independent variables (x 2 levels each), and Attribute (x 7 levels) and Speaker (x 6 levels) as within-subject independent variables.

Beliefs about rLIKE: Qualitative Analysis

Native Speaker Beliefs

A total of 54 native speakers of English (37 females, 16 males, one person identified as non-binary agender) participated in the attitudinal survey administered at the end of the second part of the Qualtrics questionnaire.

Thirteen out of fifty-four NSs who participated in the survey were able to identify rLIKE as the phenomenon the overall study was targeting. Most of these NSs also mentioned other elements of the vernacular, such as *um* and *and stuff* and described rLIKE and these elements as "filler words", "common words/phrases", "natural things", "words/phrases such as", or even "slang". When asked about their thoughts on rLIKE directly, many other NSs used the term "filler words" as well; the participants also frequently used this term when reporting other

people's attitudes towards rLIKE (e.g., comments from parents or teachers, articles seen in mass media, etc.).

Remarkably many (44) participants reported the social saliency of the phenomenon and cited the comments on rLIKE usage they heard from various people around them. Many such comments (18) were reported to have been made by respondents' school teachers, college professors or public speaking instructors. Twenty-seven participants had been called out for using rLIKE themselves, mainly by parents (7), other family members, such as grandparents, siblings or cousins (5), or teachers/professors (4). Several people told stories about family members counting rLIKEs in their speech or teachers enforcing "no like" rules in the classroom or inventing other ways of reducing rLIKE usage by the students: "One teacher in high school would make us do this thing where we had to get up in front of the class and just talk about something … for as long as we could without saying "like", "uh", "um"".

Table 5.1
Frequency of LIKE usage as reported by NSs

	All the time	Frequently	Sometimes	Hardly ever	Never
NSs themselves	4	27	19	3	1
Other people of the same age	15	30	3	1	0

As for the frequency of usage, participants were generally more likely to report the high frequency of rLIKE usage by other people rather than themselves (see Table 5.1).

Twelve people mentioned they were actively working on either eliminating rLIKE from their repertoire entirely or significantly reducing its frequency, which was often prompted by other people's comments or even by real-life consequences of rLIKE usage, such as lower grades. For example, Swaswan (m, 19) wrote in a clearly self-mocking manner: "In like 8th grade I gave a presentation and lost points because I like used it too much". Some participants felt that comments on the usage of rLIKE were intimidating or, as one participant stated,

misogynistic, because rLIKE can be seen as "a traditionally female speech pattern that is often weaponized against women" (Emily, f, 19). Only a few people mentioned seeing comments on rLIKE in mass media or social media, but such comments were typically described as stereotyping and even mocking blondes, "valley girls", teenagers, or generally the younger generation with which the participants identify: "Many articles trashing on Millennials and other young people will point out how often we use the word 'like' as a way to make us seem unintelligent, like this generation is failing the country or something" (Allie, f, 20). It is important to note that the emphasis on "us" (young people) vs. "them" (older people and/or people of authority) is not unique to this participant. For example, Biggy Smalls (m, 19) wrote: "Normally it's just parents mocking how much we use it [rLIKE]"; several other NSs used similar language to establish identification with "us" (young people who use rLIKE) and report that "they" (older people, parents or teachers) would criticize the language of the youth.

When asked what kind of people use rLIKE most frequently, the participants overwhelmingly pointed to younger people (20 mentions per 25 responses), with some specifically mentioning teenagers, adolescents, high school students, or even the age range (10-33). Fifteen participants mentioned gender as a factor, fourteen of them pointing to females and one to "gay guys who use gay voice". While some participants referred to such broad categories as "college women", "younger females", or "adolescent/teenage girls" others were more particular and pointed to "preppy white girls", "upper middle class white girls", "valley girls", or described rLIKE as "a white girl thing". While only three participants mentioned race at all, J-Train (m, 22) provided an interesting note when reflecting upon the other part of the study (Syntactic Judgment task, see Chapter Four): "It seems like there were a few phrases spoken that would usually be found in Ebonics. The inclusion of all the "likes" really distanced the rest of

the phrases from those containing Ebonics." Thus, rLIKE can be seen as a vernacular element not only more frequently used by white people but as a symbolically non-black phenomenon.

Only one person mentioned Californians as heavy rLIKE users, citing personal experience of growing up in Southern California where "literally everyone" uses it, and two people found it difficult to point to any social group in particular.

Thirty-nine people stated that they avoid or try to avoid using rLIKE in "any kind of professional setting/situation". Specific examples of such settings and situations included a job interview, public speaking, giving a presentation (including presentations in the classroom), etc. The majority of the participants also reported avoiding rLIKE around certain people, such as teachers/professors, bosses (six mentions per each category), "people of authority" (four mentions), professionals/colleagues/peers (three mentions), as well as "older people", "anyone that's not a peer", "smart people", "those who say "like" a lot [themselves]", parents or adults in general.

Eight people admitted that they were avoiding rLIKE in order to sound professional (or "not to sound unprofessional"); other reasons included the desire to sound sophisticated, to sound older, to impress someone, or to not sound uncertain, uneducated, unintelligent, or stupid. The reasons to use rLIKE more often included telling stories (which is "difficult to do without like"), describing something, being around friends (especially if they use rLIKE as well), or being nervous. Interestingly, while some participants simply reported that they would use rLIKE more frequently when nervous, others noted that overusing rLIKE may reveal someone's nervousness to other people.

While the survey did not include any questions directly asking what impression the presence of rLIKE produces, or, in other words, how rLIKE users may be perceived by others,

the participants made multiple (twenty-nine, to be precise) comments on that matter answering other questions in the survey. For example, skittles (m, 19), answering the question “Are there situations in which you are deliberately trying not to use “like”?”, wrote that he tries not to use LIKE when giving presentations because “it sounds uneducated and unprofessional”. The comments were surprisingly uniform, and three common types of impression emerged: using rLIKE makes people sound less intelligent, less educated, and lacking certainty about what they are talking about (see participant comments in Table 5.2).

Only one participant wrote about the “positive” side of rLIKE and described it as a rapport-establishing device: “I think it just sounds more informal and nonthreatening when using it [rLIKE] in front of an adjective and makes other people relate more to you in a way. It's just comfortable to say” (Skittles, m, 19). Another participant indirectly confirmed it by describing the opposite effect, when the deliberate avoidance of rLIKE created distance between peers and provoked negative emotions: “I think annoying people who think they're really smart purposefully try to use it less than everybody else” (Kate, f, 20).

Table 5.2

NS Statements on the Association Between rLIKE and Lack of Intelligence, Education and Situational Confidence

	Uneducated	Unintelligent	Uncertain/Unsure/Unprepared
Participant statements	• it makes you sound less educated • it sounds uneducated and unprofessional • it can sometimes make us sound like we're uneducated • I have heard professors and teachers talk about how "uneducated" it sounds to say the word "like" so many times in a phrase or sentence • I don't think it sounds educated • it makes you sound really uneducated ... • [I was told] it sounded uneducated ... • using the word "like" gives off an uneducated or unprepared vibe	• it makes me sound less intelligent • it makes me sound more intelligent if I don't use the word "like" • I did have one teacher mention to my high school English class how unintelligent it makes us sound • [My Dad] said it made me sound unintelligent ... • I think the main opinion out there is that people who say "like" a lot are dumb • they [teachers] said it made me sound silly and less intelligent • it makes you sound really uneducated and ditsy • I don't want to sound stupid [which is why I try to avoid "like"]	• it makes me sound uncertain of what I'm trying to say • it makes a person seem like they don't know what they're talking about • it gives off a vibe that I'm nervous or am unsure about what I am talking about • it makes me sound unsure or nervous about the situation and comes off as if I don't care as much • [My Dad] said it made me sound unintelligent and unsure of what I was saying, like I was searching for a meaning I couldn't find • [They said] it sounded uneducated and that you didn't know what you were talking about • using the word "like" gives off an uneducated or unprepared vibe

Non-Native Speakers

Non-native-speaking participants did not complete the attitudinal survey; instead, they were interviewed at the end of the second session of the study. Twenty-five NNSs (15 females, 10 males) took part in the interview. The interview served several purposes: debriefing participants about the purpose of the study, asking direct questions about their attitudes towards rLIKE, and obtaining a second speech sample. While the questions I asked during the interview were essentially the same as presented to the NSs in the attitudinal survey, the interview format did not allow for the same level of concreteness in participant responses. Some participants would digress, others would use any question as an opportunity to talk about their views of the English language in general. I chose the strategy of not pressing for a direct answer as it could potentially force the participants to stop responding altogether, which is why the attitudinal data received from the NNSs are at the same time richer and less precise or quantifiable than the data provided by NSs. In order to capture this impression, as well as to preserve all "likes" in the discussion of rLIKE, I chose to use as much direct speech as possible when writing this part of the chapter, instead of reformulating and summarizing, so that the voices of the participants could be heard, and their sharp observation skills and linguistic intuitions appreciated.

General Impressions. Each interview began with an inquiry into whether the participants were able to figure out the purpose of the study (only four NNSs mentioned noticing rLIKE; cf the 24% of NS participants who figured it out), then proceeded to debriefing. The term "remarkable LIKE" was not mentioned, but I described it as "the word 'like' which is not used as a verb, but conversationally/colloquially, in various places in a sentence" and often gave several examples of rLIKE (as a discourse marker, a discourse particle, and a quotative). Most of the participants immediately (even before I gave any examples) provided a verbal or nonverbal

acknowledgment that they were aware of the phenomenon and frequently notice it in speech. Below are three excerpts from the interviews which all illustrate the instant recognition of the phenomenon.

1) IZ = Irina Zaykovskaya, A = Aisya, f, 22, Malaysia.

IZ: What I'm actually looking at is how people use and perceive the word “like” in English. Not as a verb, but the colloquial way it's used. You're nodding, so you know what I mean?

A: Yeah, because there's a lot of “like”, “like”, “like”, there's American things.

IZ: American things?

A: Yeah.

IZ: So, why is it American?

A: I dunno, because... They like... They love to use “like”. And I learn from that too. I “like, like, like”... And that's, uh, like that makes our speech more understandable and more people will understand more.

2) IZ = Irina Zaykovskaya, J = John, m, 29, Brazil.

IZ: I'm looking at how people use and perceive the word "like".

J: Oh, yeah. I noticed a lot of “likes” in the...

I: And you know which “like” I'm talking about, not the verb to like something, but the colloquial like.

J: Yeah, the vicious[11].

3) IZ = Irina Zaykovskaya, M = Marsha, f, 21, Indonesia.

[11] “Vicious words” is the term John invented for frequent vernacular elements he does not approve of, such as *um, uh,* and *like*.

IZ: I'm looking at how people use and perceive, react to the word "like" in the English language. I obviously don't mean the verb.

M: Yeah. They always say “like”.

IZ: So, did you notice it in the recordings?

M: Mm hmm. I noticed it a lot in the second interview and... Um, I thought it was very familiar, just like that. And that's sort of when you know how much people say the word “like” and everything in America.

Only two participants requested further elaboration and examples after I mentioned rLIKE for the first time, and eventually both demonstrated understanding by giving relevant examples themselves. All of the NNSs subsequently admitted using rLIKE in their speech, most of them frequently, or “a lot”, (12 participants) or sometimes (10 participants). Indeed, all of them used rLIKE during both the short sociolinguistic and attitudinal interviews, although with highly varying frequency (see Chapter Four). Two of them specified that they use rLIKE only “in casual conversation”, another five reported successful or ongoing attempts to reduce the frequency of rLIKE in their speech. One NNS mentioned “you know” as his “preferred filler word”; several other NNSs did not discuss “you know” but used it more frequently than rLIKE during the interview.

Several participants mentioned the equivalents of rLIKE in their native language or other languages they speak, such as Swahili, Thai, or Japanese[12]. For example, Su (f, 23, Kenya) not only recalled an equivalent of rLIKE in Swahili, but discussed her mother’s attitude towards it, which is strikingly similar to the accounts provided by native speakers of English whose parents do not approve of rLIKE usage:

[12] Unfortunately, very little research has been done on rLIKE equivalents in other languages, especially non-European. Because of that, I only report what the participants told me, without corroboration from a reliable source.

IZ: Do you think you used “like” before coming here?

S: Yeah. Yeah, we do. I mean, it doesn't have to be the English “like”, but in Swahili. Yeah, we'll still use “like” a lot.

IZ: You mean the word itself or an analogue?

S: Yeah. Like, um... “Like” in Swahili. You know, like, how I can say “book” and then I can say “book” in Swahili, yeah. Yeah, so, like, how does it sound. So, it's “mfano”[13] in Swahili, so we'll use “mfano” a lot, so much, yeah.

IZ: And you use it pretty much the same way?

S: Yeah, and my mom doesn't like it when I use it much... When I use it a lot, 'cause she's like, it's...you're not...it's...you are...it's like, you're not confident in what you're saying, so don't use that word. So when I speak to other people, I usually don't mind, but when I'm speaking to my mom, I'm like very conscious, 'cause she'll be like, ‘Don't use that word,’ she'll say, ‘Don't use it, it's depicting that you're not confident in what you're saying or, you know, what you're talking about, so speak straightforward.’ Yeah. So when I speak to my mom, I try not to use it, but with other people it's hard to avoid it.

Frequency of Usage and Acquisition. Twenty-three (out of 25) participants reported that people around them used rLIKE “a lot” or “always”. However, only a few people specified the social profile of rLIKE users beyond “all Americans”, and these specifications were much less uniform than those of NSs. Furthermore, most of the NNSs appeared to be relying on their own observations rather than on any kind of knowledge received from the native speakers directly (e.g., from the reading materials or in the form of someone’s comments). In fact, very few people reported having heard any comments on rLIKE, and most of those (three out of five)

[13] The direct English translation of “mfano” is “example”; “mfano” is a colloquially used short form of “kwa mfano” (“for example”).

happened to belong to the same public speaking club whose members were supposed to eliminate rLIKE from their speech: “OK, so I'm in this club called Toastmasters, and so we usually count “uh” and “like”, and they talk about like finding like filler words instead of those words, so like that's the only place...” (Juju, f, 22, Nigeria). Observing the speech of native speakers and adjusting one’s own behavior and/or forming certain opinions, however, was a distinct pattern that emerged from the interview data.

Thus, answering my question about how she got to use rLIKE, Aisya (f, 22, Malaysia) said: “Because, as I've told you before, I love to observe people and so... Why... I realized they use “like” a lot. “Like”, “like”. So I tried to use it and it works well.” The words “tried to use it” implied a conscious decision, and when I asked if it was so, Aisya confirmed and added that it too some practice to introduce this feature into her speech. CY (m, 21, China) told a similar story, underscoring that his conscious decision to use rLIKE was motivated by the desire to not just be understood by the native speakers but to sound natural:

> C: “I was like”, this phrase - I began to use this phrase after I came to the United States.
>
> IZ: Do you remember how soon? Was it a conscious decision? Were you following the native speakers when it somehow happened?
>
> C: Yes. I was following the native speakers, because they were like, ‘I was like.’
>
> IZ: But was it a conscious decision? Was it “because they're doing it, I should, too”?
>
> C: Yeah, it was a conscious decision. Because I... I knew that... What I learned from our school system was right but not proper or not natural. Not natural.

This last remark is significant and telling. While CY’s phrasing (“right but not proper”) is the most concise, several other participants also mentioned the mismatch between the English they learned in the classroom in their home country and the English they heard around them in

the United States. Thus, Griggs (m, 22, China) said that his English teachers in China spoke "Changlish", which he described not as broken English but as "adamant to the grammars" and "noncolloquial".

rLIKE as an American Phenomenon. Most of the participants noted that even if they had heard rLIKE before coming to the U.S. at all, it was in the speech of people who spent extended time in America. Emily's (f, 23, Thailand) story is just one of the many: "I have a friend who... She, like her... how do you say? It's like, she's came to study here, in the U.S., for a year, exchange student. And when, you know, she got back to Thailand, she used "like" a lot, because she was with American girls." As Iris (f, 20, China), summarized, "if I hear somebody, like not American students but international students, if I hear them say "like, like", so I know this person has been here for a while."

The idea of rLIKE being an "American thing", as already evident from the multiple quotes above, was shared by many NNSs. Below is the excerpt from the interview with Danni (f, 22, Thailand) who went to an international school with an American curriculum back in Thailand:

> D: So, I never took TOEFL, because it's computer-based and I'm not good at computer ... but I took the IELTS, that's the British one .... And during the interview part, that's one-to-one... And then... So, basically they have like British people and after the interview he's like, 'You're like a really American curriculum student, because you speak "like" a lot.' I'm like, 'I do? I did?' I didn't even notice it, I didn't know it, yeah.
>
> IZ: So that was even before you came here?
>
> D: Yeah. So, because, say, I went to American curriculum school, every media, entertainment or like the way I grew up, culture and other things kind of tend to

American. So when YouTube, I watch American YouTube, I don't watch Thai YouTube. Movie, music, songs, everything, like both American way. Yeah.

I: So you think it's an American thing?

D: Kind of, yeah. British people don't speak “like, like, like”. It's just like a word that you didn't know it, like it's like a word to link the sentence when you didn't know something, what to say. Or what to describe something. Yeah. What do British people do? And say... I dunno. Yeah, I dunno.

Danni told me that American culture constituted a very important part of her life even before she came to the United States. In fact, Danni was one of the very few NNSs who was not only familiar with rLIKE as a discourse feature before arrival but also used it themselves. Furthermore, she seemed to have been oblivious to her usage until a British IELTS examiner told her that, so his comment served as a revelation. She even concluded from it that rLIKE was used only by Americans, even though her linguistic intuition suggested that there must be a word with similar functions in the British English, as evident from her pondering of what the British people do or say when they needed to describe something or buy themselves some time during conversation. Ibra (m, 22, The Gambia) reported that rLIKE being an American phenomenon was not the only but the first idea that came up when he was discussing it with his friends:

I: Because we didn't know how did it start or why are people using it, 'cause, I mean... Like, you know, like, like, where does it come from? I personally don't understand, you know, I don't really know where to find out about “like”. So we were just thinking. So we're like, oh, it's a U.S., American thing, you know? That's what they were saying. 'Cause you won't find it in any other country, you know? Like, like, like... But somewhere like, ooh, but then, even in Great Britain, so maybe it's like something in

English-speaking countries. You know, some people had... Had different perspectives about them, some were like, oh, maybe movies. But which movie and how and how did it go, so that everyone's using it. I was like, I don't know. I don't know. I don't know. So, at the end it just ended up, like everyone was like, no one knows, so it just ended like that."

It is evident from Ibra's phrasing ("even in Great Britain") that, while he acknowledges that rLIKE is English-speaking countries, it is still more strongly associated with the U.S. and American English. Su (f, 23, Kenya) said that "definitely Americans use it more [than other people]", Zhang (m, 21, China) called rLIKE "the case of really really American English", An (f, 21, Taiwan), responding to my question ("Is it associated with any particular type of people?"), simply answered: "Americans". I also quoted Marsha (f, 21, Indonesia) and Aisya (f, 22, Malaysia) above; Zanah (f, 22, Malaysia), Emily (f, 23, Thailand), Maggie (f, 21, China) and Ryan (m, 23, Vietnam) gave similar responses.

rLIKE in Non-Native Speech. Time spent in the U.S., exposure to native speech, and English proficiency/fluency are the three factors that, according to the participants of my study, predict whether a non-native speaker of English would use rLIKE. MT (m, 21, China) considered proficiency to be the main factor: "For example, the Indian students, they use it, because they speak it well, English, than Chinese student. Yeah. Good Chinese student, I mean... Chinese students who are good in English, when they speak, they will use the... They will use the "like" word." Marsha (f, 21, Indonesia) agreed: "I think the international students that have really poor or really weak English language skills don't tend to do that, but the ones that are like me and can speak English and are fluent and all, I think they tend to do that along with the Americans." Aisya (f, 22, Malaysia) elaborated, underscoring the importance of fluency: "If you are not fluent, you cannot use "like", because you will like stutter. Like, it needs fluent and you

need to speak fast to use “like”.” Iris (f, 20, China), however, suggested that contact with native speakers was crucial: “Cause some visiting scholars, they don’t… Like, they don’t interact with Americans much. Even though they have projects but they still don’t speak English much. They won’t use a lot of things. They don’t use a lot of “likes”.” Ryan (m, 23, Vietnam) supported this observation: “I think it has to do with how exposed they are to other people, 'cause those using the word “like”, they have a very wide group of people, they hang out with like multiple people from different culture and they hang out with a lot of like American friends, too. And the ones that don't speak like that, they tend to like club into their communities more often than not. So that's why we don't... They don't use the word “like” more, I think.”

rLIKE as an Identity-Building Tool. The incorporation of rLIKE in one’s speech may be seen not only as an indicator of fluency but also as a sign of the desire to assimilate, to blend in. For example, Katy (f, 21, Malaysia) told me about the way she adjusts her speech depending on what kind of English her interlocutor uses.

IZ: Do you use it [rLIKE]?

K: I use it.

IZ: Do you think you use it a lot?

K: Maybe. Especially when I talk to Americans. Like right now.

IZ: Well, I'm not American.

K: Well, I know you're not American, but just sounds like American.

IZ: So you start using it more when you hear someone speaking American English?

K: Yeah, so, back in my country, I think... I think my language is different. My English language is different. When I come here, I kinda blend in the language, so that's why I use “like” in my language here.

Thus, Katy made an important distinction between the English she uses back home, in Malaysia, and the American English she uses at Michigan State, the latter featuring rLIKE as an important element of the vernacular. Ibra (m, 22, The Gambia) even admitted that he felt as if he "had to" use rLIKE in order to blend in with the other students on campus.

I: I just hear it so much that it just registers, it just becomes something that I feel like I have to say naturally, you know. So yeah.

IZ: You feel like you have to, so it means… do you mean you are deliberately doing it or it kind of pops out and you don't control it?

I: Uh I would say at the initial stages it was deliberate.

IZ: So you were trying to use it?

I: Yeah, trying to, yeah, yeah… and then at the initial stages, you know, cause, you know, cause again, if you come into a culture and, you know, you wanna, you know, you wanna relate...

IZ: You wanna blend in?

I: Yeah, wanna blend in, so, you know, you try to use it.

Maggie (f, 21, China) also supported the idea of deliberate usage of rLIKE as a mechanism helping international students to become a part of the new culture and even used the same semi-modal verb with the meaning of obligation, "have to": "I know those specific type of Chinese, they want to, they want to get involved in American culture and they have to, um, they tend to, like, um... Uh... Copying what the American do when they're talking. So they just copy the "like" habit."

It can be concluded that international students, who find themselves immersed in the linguistic environment of an American college campus with the young native-speaking majority,

tend to consciously observe the linguistic behavior of their peers and are very sensitive to the highly frequent discourse features. Their own behavior, however, seems to depend not on whether they become aware of the frequency of rLIKE (the data I collected suggest that they all do) but on how they want to position themselves with regard to their native-speaking peers. While some of them (e.g., Aisya, Katy, Ibra, the unnamed students mentioned by Maggie) may invest into adopting the "like habit" and thus blending in, others may prefer to preserve their foreign identity. For example, Marsha (f, 21, Indonesia) confessed that she was scared that using rLIKE frequently would lead to losing her Indonesian identity when speaking English:

> M: And I would make a mental note to myself that, okay, I am going to try to not say it that often. But it failed, because I would forget. Um...
>
> IZ: But why would you want yourself to stop in the first place?
>
> M: Um, I'm scared that I would lose… like, I would lose like my Indonesian-ness, quote-unquote, I didn't want to be influenced or affected by the culture here. I wanted to be original and different, so, that's why...
>
> IZ: So do you think that makes you less Indonesian?
>
> M: Yeah. I think, I think it makes me less Indonesian and I think that's a fear of mine. To be immersed or to assimilate to this culture. Because I do value originality and individuality very much. I think being different is important for me. Yeah.

In Marsha's opinion, using rLIKE is a signal of American identity, which she, with her impeccable English and despite enjoying her studies and life on campus and being invested into her future career in American media (she was running her own podcast on the local radio at the time and planning to expand the audience), clearly did not want to develop and project to the outside world.

Stylistic Awareness. Only a few participants indicated that they were trying to eliminate rLIKE from their speech, and their reasons for that were also concerned with the image they would be projecting, but not related to the symbolic value of rLIKE as "all things American". These participants observed not only the fact that Americans tend to use rLIKE frequently but that some Americans refrain from it altogether or under certain circumstances. Below is the excerpt from the interview with Frank (m, 24, Pakistan) in which he explained why he refrains from using rLIKE:

IZ: And do you yourself use it?

F: No, I refrain from it.

IZ: You consciously refrain from it?

F: Yeah.

IZ: And was there a time when you decided to or did you just never use it?

F: I did, eventually, but, like, rarely in the beginning. Then like it would come out in my professional conversations or, you know, talking to... In interviews or whatever. So I just tried to refrain and remove that word from my vocabulary, so I'm not... I don't unintentionally use it.

IZ: Was it your own decision or did someone tell you you shouldn't use it?

F: I guess both. 'Cause I was talking to someone older than me, you know, a working professional, and they didn't seem to appreciate the word "like" a lot. That was in my freshman year. So then I was like, okay, that's not a good word to use anymore *laughs* I just...

I: Was it just your impression or did they say that you shouldn't?

F: They didn't mention anything, but I think it's just... 'Cause they weren't... I was the only one using that word. I didn't see many professional people using it around, so I just thought it was best not to use it. And then I saw all the kids using it and not any other word and I thought, okay, I don't need it anymore, then.

It is noticeable that Frank could not remember anyone telling him openly that using rLIKE would make him sound less professional, he only relied on his observation skills to realize that "professional people" did not use it, or at least did not use it frequently, while "all the kids" did. It gave him the choice of image he could project; he chose the image of the professional and thus decided to avoid rLIKE in his speech (in which he was not entirely successful, as evident from the transcription, although he used rLIKE at a frequency of 12.28 tokens per 1,000 words, which was lower than the group average of 27.65%). Carlos (m, 25, Dominican Republic) told a similar story:

C: Cause you're just repeating the same word over and over when you don't have to. And I've used it to like since I've been here for four years and I've been trying to also speak like Americans. I use it sometimes too much.

IZ: Uh huh. So do you use it consciously?

C: Unconsciously. Like I use it but then I'll be like, okay, I'm using it too much. I need [UNCLEAR]. Like right now I just need to like no, I don't have to use "like" in that sentence.

IZ: So you're trying to control your speech?

C: Yeah, that's something I want to work on, yeah, 'cause I already made that like… English speaking ways.

IZ: Uh huh. But is it a good thing or a bad thing that it is part of your English-speaking self?

C: To me it's a bad thing. I think I sound less educated when I use it too much.

IZ: Wouldn't you sound more natural? If you're saying that everybody is using it?

C: Yeah, I guess I'm saying I'm more like people from Michigan or people I know here, college students, but I don't want to sound like them.

IZ: Uh huh, you want to sound educated?

C: Yeah, I want to sound like Neil deGrasse Tyson.

IZ: Ah!

C: Or Obama, I don't know.

IZ: Uh huh.

C: Yeah.

IZ: And Obama doesn't use “like”?

C: I bet he does but when I listen to him, I don't hear like any “like” words that's supposed to be, so it doesn't become annoying because of that.

As Frank, Carlos also made a clear distinction between the way students speak (i.e., they use rLIKE frequently) and the way famous public speakers, such as Neil deGrasse Tyson or Barack Obama, speak (i.e., they do not use rLIKE at all or use it so infrequently that it is hardly noticeable). Furthermore, he too made a deliberate choice of identity he wanted to adopt and the image he wanted to project using the linguistic resources available to him: He decided that by refraining from rLIKE he would be able not to sound like a college student but to project an image of a well-spoken, educated professional. Similarly to Frank, he used rLIKE at a lower-than-average frequency of 16.4 tokens per 1,000 words).

Several other participants also acknowledged the value of rLIKE as an atmosphere creating tool and reported varying their own performance accordingly. According to Ange (f, 22, Rwanda), rLIKE makes conversation sound “informal and not serious”, and Ibra (m, 22, The Gambia) deliberately uses rLIKE “to introduce an informal situation … that informal environment”, even during an interview which “should be, uh, maybe formal or something”. Juju (f, 22, Nigeria) uses rLIKE “like a thousand times” in informal conversations with her friends but when talking to “like professors or like people in a higher rank”, she tends to “watch like every word that comes out of [her] mind”. Anna (f, 20, China) avoids rLIKE when she wants to produce a good impression on the other professionals but uses it to create casual atmosphere of mutual trust when necessary:

> A: When I want to, uh, you know, more professional setting or more academic setting, when I answer questions from, you know, answer questions from my professors, I don't... I try to avoid saying “like”. I want to be specific so that I can really answer the question, not a vague answer. Yeah, so, in a professional setting if I use “like”, the word “like” a lot, it sounds not professional. It sounds like I'm not really sure I know what I'm doing, you know. If I... Yeah, certain... Same thing, if I use specific words, um, it just, it's just people will have a better, you know, impression?
>
> IZ: Did anyone tell you that or is that your own idea?
>
> A: It's both. I... Personal feeling, like if this person uses “like” a lot, he's more casual, he's more chill, not, um...
>
> IZ: Chill is kind of a good thing. Is it a good thing?
>
> A: Well, sometimes is. It is. When you're getting to a very serious professional, then it's not a good thing to be, you know, so casual.

IZ: And when is it good?

A: It's good when, you know, when... When you want to get to know this person and they're casual, they're open, so that was a good. That's when it's good.

Overall, however, the level of stylistic awareness was low: only six NNSs acknowledged the differences in rLIKE usage in different contexts. This is consistent with Meyerhoff and Schleef's (2012) findings with regard to the usage of (ing) by Polish teenagers in Scotland who appeared to have acquired some of the linguistic constraints on -in' usage, as well as the gender constraint, but did not demonstrate any style-shifting. As the researchers noted, "we might expect the Polish teenagers to acquire or create a system of stylistic stratification of the variable before they acquire or create stratification based on gender or friendship network. And yet, they do not." (p. 411).

Qualities Associated with rLIKE. While not all participants mentioned the difference between formal and informal communication with regard to rLIKE usage, those who did reported impressions similar to those of native speakers: the impression of unprofessionalism (discussed by Frank and Anna), uneducatedness (mentioned by Carlos), and lack of confidence in or knowledge about the subject matter (brought up by Anna). Among these people, only Anna reported reading career advice articles which warned prospective interns or interviewees against overusing rLIKE; Carlos and Frank said that their opinion on the impression rLIKE users make on other people resulted solely from observation of native speaker behavior in various contexts. Several other participants also mentioned that they were previously advised against using rLIKE in writing (with a suggestion that they should use "such as" instead) or when speaking in public (e.g., giving presentations).

It can be concluded that, while the association between rLIKE and American English is clearly and strongly established for the majority or international students who speak L2 English, not many of these students are aware of the stylistic constraints on the rLIKE usage. However, the students who are aware of it tend to share some of the native attitudes towards it (e.g., the association between rLIKE usage and lack of professionalism, education, or confidence in one's knowledge). At the same time, none of the NNSs interviewed said that people who use rLIKE sound unintelligent (silly, dumb, stupid, etc.), which seems to be, along with other subjective attributes, an important component of the native beliefs about rLIKE. One participant, Ryan (m, 23, Vietnam) mentioned the word "unintelligent" with regard to rLIKE usage, but only when telling me about an argument he happened to overhear:

> R: My friend actually had an argument with another friend because of that. 'Cause, um, so, story time. So, he and another friend of mine, they're very close, too, because one of them is from Korean, from North... South Korea, himself, and the other one is American. And because the South Korean guy, he studied, um, you know, English here, and he mostly speaking... And he speaks with a lot of people and they use the word "like" a lot. And he often use it, and then, one time, my American one friend just put out, 'Hey, dude, don't... stop using the word "like", it makes you sound unintelligent.' It pissed the other guy off. Like, they got into an argument. Because, to him, using the word "like" is how he speak and my American one friend should respect that, but he did not, like they argued - oh, it makes you unintelligent, oh, you don't know what you're talking about. So there was a big fight. I think it's just a misconception or misunderstanding of their culture, or like how they come across as well.

It is notable that Ryan abstained from taking any side in this argument; he merely presented two different points of view. Thus, even post-hoc he did not support the "native" point of view (that rLIKE makes people sound unintelligent); instead, he suggested that a misunderstanding based on the cultural differences had occurred. Therefore, it cannot be assumed that Ryan shared the native perception of rLIKE as a trademark feature of less intelligent speakers.

rLIKE User Profile. Most of the NNS interviewees also did not tend to describe rLIKE as a "female thing" or a "white thing", although many mentioned young age as a prominent characteristic of rLIKE users (the descriptions of rLIKE users included the words "youth", "young adults", "teenagers", "high school students", "college students", "anyone younger than 25"). Two participants mentioned that they observed their peers use rLIKE while the professors were avoiding it, but these participants did not specify whether they perceived that difference to be motivated primarily by age or level of education and social status. Some NNSs explicitly stated that rLIKE frequency reduces as people age, with one participant suggesting rather a rapid pace: "I have conversation with freshmen, high school junior in the U.S. It seems like as they grow... Like, you go to junior, you talk to juniors and seniors, it lessens a bit. It lessens a bit. That's just my own... They don't use it as much as a freshman coming or a sophomore, you know." (Ibra, m, 22, The Gambia). Importantly, none of the participants suggested that older people may be more likely to use rLIKE.

No uniform racial profile of rLIKE users emerged from the NNS data. Four participants said that rLIKE usage was more characteristic of white people but were cautious enough to note that their observation may be biased due to white people being the majority on campus (which limits their exposure to black speakers). One participant (Zanah, f, 22, Malaysia), however,

claimed the opposite to be the case: “I think black Americans will use “like” more than white people. Because they... I dunno, maybe the slang that they're using, maybe that's why they use “like” more. I noticed that only.” At that point Zanah did not report any potential bias in her observation, but the analysis of her initial (generic) interview revealed that at the time of the interview she was employed at a canteen where all her colleagues (and thus frequent interlocutors) were African Americans.

Another interesting comment on rLIKE users with regard to race was provided by Marsha (f, 21, Indonesia). She reported hearing her non-white friends mock white women for their usage or rLIKE:

IZ: Have you ever heard anyone discuss how other people use “like”?

M: No, but I've always heard people make fun of that.

IZ: What people?

M: Usually like my friends who are not white. For some reason, there's this stereotype of mostly white girls who are the subjects of that. They are, for some reason, the subject of ridicule when it comes to saying “like” often. Yeah, um... And my friends are usually brown or black or, yeah, Indian, um, Hispanic, black...

IZ: But they still use “like”?

M: Yeah, but they still make fun of the girls.

In this conversation, Marsha was clearly trying to separate her own views and attitudes from those of her non-white friends: She used the expression “for some reason” twice, thus expressing that she could not understand where the stereotype she was reporting came from. She also acknowledged the fact that, while making fun of white girls for rLIKE usage, her non-white friends were using rLIKE themselves.

As I mentioned before, there was also no uniform gender profile of rLIKE users; only a few NNS associated rLIKE with females, while the majority claimed that “everyone”, “everyone in America”, or “all Americans” used rLIKE frequently, and a smaller part of the interviewees emphasized that rLIKE usage was more characteristic of younger speakers. However, all of the three responses that featured a specific social persona associated rLIKE with females, albeit only those possessing certain attributes.

When I asked Maggie (f, 21, China) if there is a particular type of people who use “like” more than others, she responded immediately with a detailed profile: “They have the sharp voice and clear English and they're like the most attractive girls and they just, they just come out from the high school.” Maggie could not remember anyone explicitly telling her that, so she presented this profile as a result of her own observation. Lauren (f, 22, Vietnam), on the other hand, credited Hollywood, along with her trip to California, for the social persona (a sorority girl from the West Coast) she learned to associate with rLIKE:

> L: I think girls definitely use “like” more than guys do. Um... And I know that the stereotypical California girls, sorority, the sorority type girls use a lot of “like” and I know that's, just, like, a satire stereotype that's reflected on in movie media...
>
> IZ: So you've seen it in the movies?
>
> L: Yes.
>
> IZ: And have you heard anybody talk about it? Maybe in the media, articles about “like”?
>
> L: In the media I don't, I don't really pay attention to that.
>
> IZ: Where did you get the idea that it's Californian?
>
> L: Uh... So, I think, I can't really trace back to it, but maybe it was because I was just in L.A. this past Thanksgiving, so their accent is a little bit different than the Michigan

accent and it's... It's got a twist in it, so maybe I associate that with, uh, the mainstream Hollywood movie that's set in somewhere that's warm and nice, definitely not Michigan. Uh... On the beach. And that's kind of all combined together to have the image of California, West Coast type.

In this statement, Lauren underscored how her ability to observe and compare the ways people live and speak in California and in Michigan contributed to the concept originally rooted in her knowledge of the movies in which the stereotype is perpetuated.

Ryan (m, 23, Vietnam) said that he associated rLIKE usage with “a party type”. It is interesting that initially Ryan did not specify the gender of the “party type” classmates he was talking about, but when he began to elaborate and describe specific activities such a person could be involved in, he used only the feminine pronoun:

R: I feel my classmates use it constantly and when they use it, I know... I sort of associate them with a party type and then they usually are the party type.

IZ: Uh huh! So it's the party type word?

R: I think. I think. But most of people who use that word and the way they speak, too, they use that word a lot more often than not and I associate them with like, oh, she goes to parties a lot and she drinks a lot as well. And usually they do that, that kind of activities.

It must be noted that in the Vietnamese language, pronouns are marked for number, person and the social status of the interlocutors but are not marked for gender, which means that in Vietnamese, there are no distinguishable equivalents of “he” and “she”. Therefore, there is a possibility that Ryan’s use of the feminine pronoun was not indicative of whether he perceived those party type rLIKE users as mostly feminine. However, the analysis of both interviews with

Ryan showed that he did not tend to make errors of pronominal reference (i.e., he used "he" to refer to males and "she" to refer to females at all times) and his default choice of a third person pronoun to refer to an unknown person was the masculine pronoun "he". This evidence allows me to (cautiously) treat Ryan's usage of "she" in the conversation as a signal that he associates rLIKE primarily with female party people.

Meaning of rLIKE. Many NNSs appeared to be eager to discuss the meaning and functions of LIKE during the attitudinal interviews. In most cases, their approach was surprisingly language- and discourse-oriented rather than based on social meaning or style, which is why it seems appropriate to analyze these data in this chapter as they may help provide nuance and depth to the usage analysis itself. For the purposes of this analysis, I am not constraining the discussion to rLIKE only, because, while all respondents were clearly distinguishing between the verb *to like* and other functions/meanings of LIKE, they obviously did not make a fine-grained distinction between unremarkable and remarkable LIKE and often mentioned the comparative complementizer among the other functions.

Only two participants said that they did not understand the meaning or function of LIKE at all. One of them, An (f, 20, Taiwan), confessed: "We never, like, classify it and so we kind of just listen and ignore it ... because we don't know what it means". However, most other NNSs said that LIKE can be used as a filler or a hesitation marker. The ways in which the participants would describe it varied from using the terms "filler" or "hesitation" (or less conventional terms, such as "buffer") to comparing LIKE to such discourse elements as "uh" and "um", to saying that LIKE allows the speaker to "buy time" (Ibra, m, 22, The Gambia). As Anna (f, 20, China) put it, "when they say 'like', they're thinking, they are trying to process". Maggie (f, 21, China) claimed that "it's just like, um, it's just like the function of 'ummm', 'ummm', I think". Juju (f,

22, Nigeria) admitted: “I feel like I use it when I’m trying to like think about what to say, or when I’m short of words, or when I’m trying to put my thoughts together”.

However, many NNSs also demonstrated awareness of other possible meanings/functions of LIKE. The most frequently mentioned function of LIKE (a total of nine mentions) was a linking function. For example, Jenny (f, 21, China) described LIKE as “kind of a word that ... connects my previous sentence to my next sentence”. Lauren (f, 22, Vietnam) said that “LIKE can also be used as a filler, but also to link sentences or parts of the sentence together”. Danni (f, 22, Thailand) suggested that LIKE can be used when “you want to keep the connection flowing but ... you don’t want like ‘however’, you know”.

Other frequently mentioned functions included using LIKE for comparison, for clarifying or elaborating a point, and for describing things. According to Maxwell (m, 21, Malaysia), “most of us will use this type of conversation when he wants to describe something or to tell something”. Zanah (f, 22, Malaysia) said: “The only thing I notice when people tell the word ‘like’ to compare things and also to elaborate more on what they’re telling [is] that they want to extend the conversation”. Marsha (f, 21, Indonesia) also suggested that people use LIKE “to help describe, explain and make their point clearer or make what they are saying clearer”.

Often the participants would list several functions of LIKE and, in some cases, even give examples. Below is the excerpt from the interview with Ryan (m, 23, Vietnam):

IZ: Do you think there's just one "like" or there are multiple functions of this word?

R: I think there's many functions of it. Depends on the nuances [of] how you use it.

IZ: Can you give any examples?

R: Uh... "So I was like", or "he like me, but I dunno, like". That's like "you mean like this?" For example, like that. Those are the two different ones I can give you off the top

of my head right now. The first one... It's more like she using it... The person using it as a connector. As in a connector in speaking language. But the other one is more like comparison. "Like this". So, yeah.

It is not entirely clear what Ryan meant by two different functions, because in the first sentence of his response, in addition to the verb, he gave examples of quotative LIKE, and what sounded like a clause-final like but could have been a clause-initial LIKE which was not followed by a clause. It is not possible to determine which of them he called a connector. However, his usage of all these examples shows his awareness of the multitude of LIKE functions, even though he lacks metalanguage and/or the linguistic knowledge that would help him provide more precise commentary for the examples he gave.

As none of the participants reported taking any classes in linguistics, they all predictably struggled with elaborating on the intuitions they had about the meaning of LIKE. For example, Carlos (m, 25, Dominican Republic) suggested that LIKE may be used “to describe things” but then proceeded to discuss the meaning of vagueness that LIKE may bring to an utterance.

C: But sometimes even when you're describing something, you don't need to use it. 'Cause as I said before if you think someone is nice, you don't have to say 'Oh, I think she's like nice.' No, ‘I think she's nice’. You know, that ‘like’ over there is not necessary.

IZ: Mmhm. Doesn't it add any meaning? Is there a difference between uh ‘she's nice’ and ‘she's like nice’?

C: Yeah I would say when you say she's like nice, it sounds like you're not sure.

I: Mmhm.

C: You're not sure that's the best word you can describe her. Whereas when you say ‘Well, I think she's nice’, it looks like you're more like sure about what you think about that person.

It is important to note that in this excerpt Carlos talks about potential uncertainty on the part of the speaker who is not sure whether a given word is the most suitable in a given context, which makes this statement different from multiple accounts of perception of LIKE users as unsure or uncertain that surfaced during the attitudinal interview.

John (m, 29, Brazil) also recognized multiple functions of LIKE in the following excerpt:

IZ: And what does this "like" mean? What do people use it for? Or why they use it?

J: Why they use "like"? Despite of comparison?

IZ: Yeah, apart from comparison or, obviously, the verb "to like something".

J: I think people use "like" when they're just thinking on what they're trying to say or to explain, that comes naturally for most people, actually. I think, "like" is the word that comes in mind when elaborating a thought.

IZ: So, does it help pause? Does it buy you more time or... something else?

J: No, I think it's just... it makes the bridge between parts. Even though the meaning is not quite useful in that sentences, but it kind of connects both dots of what I just said and what's coming.

In this excerpt, John demonstrated awareness of LIKE as a comparison device, as a linking device (making “bridge between parts”), and as a tool that helps elaborate a thought or explain something. Earlier in the interview, John also mentioned that LIKE could be used as a filler and called it a “linguistic vicious”, the term he coined himself for filler words.

This brief analysis shows that the level of metalinguistic awareness about the meanings and functions of LIKE is exceptionally high among the non-native-speaking participants of this study, even though none of them has background in linguistics or related areas. Furthermore, the only meanings of LIKE that would be typically discussed in English language classrooms are those of comparative complementizer or preposition. There is no apparent connection, however, between the depth of metalinguistic awareness and the frequency of rLIKE usage. For example, John, who provided multiple insights with regard to the functions and meanings of LIKE, produced only two tokens of rLIKE during the first interview (a short sociolinguistic encounter). On the other hand, Juju, who said that she uses like to fill a pause during which she would be searching for a word or putting her thoughts together, demonstrated the second highest frequency of rLIKE among all study participants (77 tokens per 1000 words). At the same time, Zhang (m, 21, China), who produced rLIKE at a similar rate of 78 tokens per 1000 words, demonstrated awareness of the multifunctionality of LIKE by saying that its function “is just connects two sentences. Maybe two related sentences. I think it can also function as "such as". Yeah... It's... It's really versatile, the word.”

Unfortunately, I did not include a separate question about the meaning of LIKE in the attitudinal survey for NSs. However, as I mentioned earlier in this chapter, NSs were asked what they thought about LIKE in general, which prompted some of them do discuss the meaning of LIKE, and in such comments LIKE was mostly described as a filler word (not necessarily by this term) and a word that does not have any meaning in particular. This is consistent with Fox Tree's (2007) finding that native speakers of English are generally not able to describe the meaning of rLIKE, even though, when presented with examples of usage, they are able to say whether LIKE

is appropriate in a given context. Thus, NNSs demonstrated a higher level of metalinguistic awareness with regard to rLIKE than NSs.

Conclusion

Native speaker beliefs about rLIKE appeared to be consistent with the popular beliefs about it, which, according to D'Arcy (2017), are shared by the majority of native speaker of English in North America:

- rLIKE is recognized as a popular cultural phenomenon associated with the Valley Girl social persona, and NSs in my study discussed various stereotypical portrayals (e.g., on TV) of blonde Californian girls mocked for their overuse of rLIKE;
- rLIKE is a feature of youth's speech, and NSs overwhelmingly talked about how rLIKE created the "us vs. them" dynamic between themselves and older people as well as discussed how older people criticize the youth for rLIKE overuse;
- rLIKE is associated with females, and the majority of NSs mentioned girls and young women as the primary users of RLIKE;
- rLIKE is a symbol of inarticulateness, and NSs remembered being coached not to use rLIKE when speaking publicly and talked about how rLIKE projects an image of a speaker who does not know what to say.

In addition to these beliefs, NSs seemed to recognize rLIKE as a predominantly white phenomenon, which aligns with Bucholz's (2011) note on "preppy whiteness" associated with rLIKE. Another important theme was that of rLIKE being unprofessional; multiple NSs reported their attempts to avoid rLIKE altogether in order to be perceived as professionals or at least reduce rLIKE usage when talking to people of authority (e.g., bosses or college professors).

While all NNSs were aware of rLIKE as a discourse phenomenon and did not need any explanation beyond the word “like” used not as a verb but conversationally in order to recognize it and become eager to discuss it, the actual beliefs about rLIKE they demonstrated differed from the native ones substantially. The most prominent theme that emerged from NNS data was the association between rLIKE and American people and American English. In other words, for NNSs in this sample, rLIKE was an overwhelmingly “American thing”. Most NNSs reported that their first encounter with rLIKE happened when they arrived in the United States or when they met a speaker of American English (a native speaker or a non-native speaker who lived in the U.S. for some time).

The association between rLIKE and Americans was much stronger than any finer distinction, although many NNSs talked about rLIKE being more characteristic of younger speakers and some NNSs mentioned females as heavier rLIKE users. Given the strength of this association, it was not surprising that many NSSs reported a conscious decision to incorporate rLIKE in their speech as a way to develop their L2 identity and to “blend in” with the native speakers. A possible link between English fluency as well as exposure to naturally occurring English speech (including the choice to socialize primarily with native or non-native speakers) was also brought up by several NNSs.

While some NNSs demonstrated awareness of the native tendency to avoid rLIKE in professional settings and their desire to avoid rLIKE to sound more professional, they also recognized the atmosphere-building power of rLIKE and discussed how the usage of rLIKE can make a positive contribution even in a professional setting, which NSs in my sample did not discuss at all, with the exception of one person who talked about rLIKE making speech non-threatening in general.

Finally, an important point of difference between native and non-native beliefs about rLIKE appears to be the perceived source of these beliefs. The overwhelming majority of NSs mentioned the influence of parents and other family members and teachers/college professors on their beliefs. Furthermore, they seemed to mostly accept the statements such as "like makes people sound dumb" made by those people at face value, without any attempt of critical reflection. NNSs, on the other hand, usually attributed their awareness of rLIKE (as well as its meaning and appropriateness in various contexts) to their own observation skills rather than to any external influence. In other words, they perceived rLIKE as an identity-building tool and claimed agency over understanding and using it.

Attitudes towards rLIKE: Experimental Data

The first part of the matched-guise task required the participants to guess the age of the speakers and freely profile them. Twenty-eight native and 26 non-native speakers of English completed the task, but the quality of the resulting native and non-native profiles differed substantially. Before presenting the resultant speaker profiles, however, I will first discuss the ways in which NSs and NNSs approached the task and what kind of social information they focused on when creating speaker profiles.

The NSs performed very well at this task: Most of the profiles they created were rich and contained various types of information. NSs tried to guess the following characteristics of each speaker:

- *Race and gender*: "This person is around 21 years old, white male" (Chris, f, 21, about Speaker 2); "I think she is an African American" (Christian, m, 18, about Speaker 4);

- *Social class*, often with a geographic or other specification: “I would assume he’s middle to upper class” (Andrey, m, 18, about Speaker 2); “23 year old male who grew up in a middle class family in a suburb in New York city” (b.p., f, 19, about Speaker 2); “middle class, suburban” (PW, f, 18, about Speaker 3);
- *Level of education*, often coupled with the age estimate: “She sounds like she attends a very prestigious college” (Al, m, 18, about Speaker 1); “ambiguous but probably high school age” (Sero, a, 19, about Speaker 6); “newly graduate” (Nox, f, 20, about Speaker 2);
- *Native speaker status*: “Comes out native but has trouble thinking of the words/what to say” (Tchaikovsky, m, 18, about Speaker 5); “Possibly not a native speaker. No accent but difficulty choosing words” (Lena, f, 19, about Speaker 5);
- *Subjective attributes*: “Can tell she is very intelligent and sure of herself” (Estelle, f, 20, about Speaker 4); “simple minded, boring, unintelligent, immature” (Jannet, f, 20, about Speaker 6);
- *Situational characteristics of the speaker or characteristics of the situation*: “sounds nervous” (Kate, f, 20, about Speaker 5); “She is probably teaching someone English” (zdravo77, m, 20, about Speaker 1); “he is probably in an interview or something” (zdravo77, m, 20, about Speaker 5);
- *Physical appearance*: “20, male, skinny, dark hair” (turtle, f, 18, about Speaker 5); “probably blonde with blue eyes” (b.p., f, 19, about Speaker 3);
- *Accent/regional affiliation*: “Probably from the Midwest because I think she said ‘pop’ for soda” (Kate, f, 20, about Speaker 3); “She is probably from Ohio, lol” (Al, m, 18, about Speaker 3); “Probably an American from the south (although

not deep south) or a little further west" (Lena, f, 19, about Speaker 3); "sounds like they grew up in the Midwest(ish) (there's no regional accent)" (Allie, f, 20, about Speaker 1);

- *Voice*: "has vocal fry" (Allie, f, 20, about Speaker 1); "voice sounds 18ish to me (not meant as a jab bro, you got the voice of a rock star)" (Tchaikovsky, m, 18, about Speaker 5);
- *Manner of speech*: "speaks quickly" (Beyoncé, f, 19, about Speaker 3); "stuttered a little" (lucy5, f, 19, about Speaker 5); "keeps repeating the same information in differently formatted sentences" (anonymous, f, 18, about Speaker 1);
- *Linguistic choices*: "the words she is using are much more in a younger vernacular" (Ramirah, f, 20, about Speaker 4); "says like a lot, lots of pauses, says um and yeah a lot" (Beyoncé, f, 19, about Speaker 1).

Certain types of social personae were mentioned ("sorority girl", "nerd", "jock"), and some NSs provided very detailed profiles of the speakers or even attempted to reconstruct their family history. Below are two examples of such detailed profiles:

1) Al (m, 18) about Speaker 6:

Sounds like a middle aged white girl. She sounds like she had a very good childhood, very sheltered. Probably lived in the suburbs and her dad was a lawyer and was secretly having an affair on her mom and everyone knew it, but no one ever brought it up.

2) b.p. (f, 19) about Speaker 2:

23 year old male who grew up in a middle class family in a suburb in New York city. His name is either David, Chad or Nick and he has never read a full book assigned to him in an English class. He is 5'12".

Judging by the length and quality of the NSs' responses, the participants were interested and invested in the task and even enjoyed it (some mentioned it in the space provided for notes at the end of the entire questionnaire). NNSs, however, appeared to have found the task difficult or misunderstood its goal altogether. About half of the NNSs treated it as a listening task and wrote brief summaries of the stories they heard or even managed to almost transcribe those stories. Some only gave the age estimate and were reluctant to volunteer any judgment of people they did not personally know, since they told me it would not be polite to do so.

Only ten NNSs (out of 26) consistently gave the kind of detailed responses that the NSs provided (e.g., "college freshman, Caucasian, probably interested in Greek life"; or "20-25, college student, white, born in the States"). The range of the types of information they attempted to provide was narrower than that of the NSs. For example, while NNSs also frequently tried to guess the speakers' gender, race, level of education, native speaker status and subjective attributes, few of the NNSs mentioned social class. There were only singular instances of comments on physical appearance, situational characteristics of the speaker, or accent/regional affiliation, as well as of social personae mentioned (three of those emerged, "party type", "frat guy" and "sorority girl"). Only one NNS (John, m, 29, Brazil) consistently mentioned the manner of speech or commented on the linguistic choices of the speakers (e.g., "struggling with word choice" or "she shows some 'linguistic vicious' as saying 'so' every end of the sentence, really speaks outwards, imposing her ideas").

NNSs' Reflection upon Judgments

In order to obtain more detailed profiles from the NSs, I asked each of them to listen to two of the recordings (chosen randomly) again and talk me through their thought process during profiling. This took place at the end of the second session of the study, just before the attitudinal

interview. By that point, the participants had completed the second part of the matched-guise task as well, so they talked mostly about the characteristics from the list that had been given to them to rate each speaker (friendly, intelligent, educated, polite, ambitious, attractive, confident). Therefore, these data could not be used as additions to the free profiling information, but instead provided important insights into the differences between the way NSs and NNSs might view certain characteristics and what they might use as sufficient reasons to perceive speakers as friendly, polite, etc.

When profiling speakers, NSs relied mostly on the form rather than on the meaning, that is, on the manner of speech and lexical or other characteristics of speech rather than on the content. Of course, I could not directly ask them whether this was the case, but there is ample circumstantial evidence thereof: NSs often mentioned speakers' voices, manner of speech of linguistic choices in order to justify the characteristics they would give to the speakers but they almost never mentioned the content (with one exception: Speaker 3 talked about moving to college and most of the participants pointed to this fact and stated that the speaker was a freshman). Also, the length and the quality of the profiles did not depend on the type of content (some Speakers were discussing the difference between certain words while others talked about themselves thus providing some factual information that could be potentially used for profiling). NNSs, on the other hand, relied on the content heavily (at least initially), which resulted in more detailed profiles of the speakers who talked about themselves. Furthermore, some NNSs provided explanation for their ratings of specific speakers which suggested that their interpretation of friendliness may be specific and based on their own experiences.

Thus, two NNSs described Speaker 1 as friendly "because she speaks slowly". This reveals at least one dimension of friendliness that is specific to non-native experience. Naturally,

it is easier to comprehend L2 speech when the speaker is not talking too fast, which is why slow pace of speaking may be perceived as a signal of consideration of the interlocutor's needs on the part of the native speaker. Others interpreted slow pace as a sign of being well-educated. Several other NNSs told me that they perceive any non-aggressive attitude as friendly; Zhang (m, 21, China) said that he perceives others "friendly if they aren't being rude" and CY (m, 21, China) treated "not showing any attitude" as a signal of friendliness. Also, many NNSs did not distinguish between politeness and friendliness. As Griggs (m, 22, China) stated, "if a person is friendly, she is polite. Can't be polite and not friendly". Other unexpected interpretations included confidence and ambitiousness. Thus, Iris (f, 20, China) said that Speaker 5 did not seem very confident to her because it did not "sound like he's bragging", and CY noted that "ambitious is not a good concept in Chinese culture" which was why he tended not to rate speakers whom he liked as ambitious. It is interesting that all of the NNSs who expressed these evidently non-native perceptions of the attributes in questions were from China, and five out of eight participants from China who participated volunteered these perceptions. However, students from China were the only large monoethnic group within my sample (the second largest group, students from Malaysia, was two times smaller), so it does not seem possible to determine whether these perceptions are characteristic of Chinese students specifically or generally of non-native-speaking international students.

Speaker Profiles

In order to create composite profiles of each speaker created by NSs and NNSs, I looked for the features/characteristics mentioned in responses both rLIKE and bare guises and not contradicted by any participant. For example, if several participants profiled a particular speaker as a white woman across both guises and none of the other participants mentioned any other race

or gender as a possibility, the composite profile of this speaker would be based on this information. Also, the "native" profiles appeared to be more reliable due to the fact that virtually all the responses included profiling, while less than half of the "non-native" profiles included data that could be used for creating composite profiles.

The profiles that emerged are summarized in Table 5.3. As can be seen, most of the profiles are quite generic, especially those created by NNS, although the composite profiles of the male speakers (Speaker 2 and Speaker 5) created by NSs were elaborate and even included descriptions of physical appearance.

This might indicate that the presence or absence of rLIKE in male and female speech may be perceived differently, although the reasons for this difference cannot be established with a sufficient degree of certainty. One possible explanation would be that, as rLIKE is strongly associated with women but not with men (according to both previous research and the attitudinal interview data in this study), listeners simply do not expect to hear rLIKE in male speech and do not "check" for its presence before making a judgment. When hearing a young woman's voice, the same listeners would be more likely to categorize the speaker as using or not using rLIKE.

Table 5.3

Composite Sociolinguistic Profiles of Speakers 1 – 6 Created by NSs and NNSs of English

Speaker	NSs	NNSs
Speaker 1 ("A cap and a hat")	White woman in her late teens, college student ✓ possibly from the Midwest (the only common guess across guises) ✓ possibly a native speaker (the only common guess across guises)	White American woman aged 22-25, college student or graduate
Speaker 2 ("A townhouse and a duplex")	Tall dark-haired white man aged 22-25, college student or graduate, who uses a lot of fillers when speaking ✓ possibly from the Midwest (the only common guess across guises) ✓ possibly a native speaker (the only common guess across guises)	White American man aged 24-27, college student or graduate, belongs to middle class or higher
Speaker 3 ("Parting with the family")	White woman aged 17-20, college student* ✓ possibly from the Midwest (the only common guess across guises)	White American woman in her late teens, college student, outgoing
Speaker 4 ("Future career")	Woman aged 26-28, college student*	Woman around 20 y.o., college student* or graduate, a good student
Speaker 5 ("Stage lighting")	Skinny and dark-haired man aged 19-24, unsure of himself, nervous during the interview and stuttering ✓ possibly white (the only common guess across guises)	White man aged 22-23, college student, insecure, unsure of himself and nervous
Speaker 6 ("Haircut")	White woman aged 20-25	White woman in her late teens, high school or college student

* Information could be inferred from the content of the story and thus does not necessarily reflect listener perceptions.

The following step in the analysis of speaker profiles, namely the by-guise analysis, could be reliably performed only for the NS data due to scarce information provided by NNSs. However, I took a note of any interesting comments by NNSs or of the comments overlapping with those given by the NSs. It must also be noted that additional free profiling data were collected from 36 native speakers of English (most of whom participated in the attitudinal questionnaire as well). Due to a technical error, the number of the participants who received different packages was not balanced and they did not perform the second part of the matched-guise experiment, so the data from this group were not used for creating composite profiles in order to preserve methodological consistency across the entire task. However, as these participants did not differ from the main NS group demographically or otherwise, I decided to add some information they provided to the analysis.

Speaker 1. While the composite profile of Speaker 1 was not very specific (a possibly Midwestern young woman, college student or graduate), the profile of the rLIKE guise was more pronounced. First, ten NSs (a total from both main and additional samples) mentioned the fact that she used rLIKE (as well as other discourse phenomena) in her speech, for example, "says like a lot, lots of pauses, says um and yeah a lot" (Beyoncé Pad Thai, f, 19). NNSs did not mention it in the written profiles but some discussed it during the subsequent reflection session; Danni (f, 19, Thailand) said that "she didn't use speaking words, she used communication words". None of the NS or NNS listeners who heard the bare guise mentioned those "communication words" (*um, yeah*) or pauses, even though they were still present. Interestingly, one NNS (lizardgod, f, 20) noticed the absence of rLIKE: "She doesn't really seem to use the word 'like' at all, which makes me think she probably isn't a teenager or maybe is just well educated". Interestingly, however, the mean age estimates for both guises given by NSs were

exactly the same, 19.8, and the NNSs though that the woman was younger in the bare guise (22.23 vs. 24.7).

Among the various other characteristics given by NSs and NNSs to Speaker 1 in her like guise, an image of a young blonde woman, privileged and possibly belonging to a sorority, emerged. Most of the participants discussed the level of her confidence, but did not agree on it, with many perceiving her as very confident and some claiming the opposite (and mentioning rLIKE usage as the reason behind their judgment). The bare guise image was less precise, not as wholesome, but more serious, with one NNS suggesting she could be an overthinker and some NSs speculating along the lines of "she works in some sort of field where she needs to analyze things in depth, perhaps journalism" (Nix, f, 20). However, for one of the NNSs, Maggie (f, 21, China), Speaker 1 even in her bare guise sounded like a "mean girl" and "traditional American": "clear voice, white woman, blonde hair, blue or green eyes, I may not like her".

Speaker 2. In his like guise, Speaker 2 was profiled as a "frat guy" or "frat bro" (four NS, one NNS mention), with some other descriptions projecting a similar image (i.e., a college student who frequented parties) even without the word "fraternity" in it. For example, Al (m, 18) described this speaker as "a white guy who graduated Florida State University a few years ago. He never took school seriously and he was a complete tool who took advantage of every girl he met at a party". The word "laid-back" was also frequently mentioned; the word "jock" was mentioned not very often across both guises. Three NSs noticed his usage of rLIKE and/or "other fillers" (*um, yeah*); one NS even hallucinated rLIKE in the bare guise. In the bare guise, Speaker 2 was mostly perceived as more serious and professional, but no distinct image of another social persona emerged. Both NSs and NNSs though that Speaker 2 was younger in his like guise (22.4 vs. 24.5 for NSs, 24.06 vs. 26.12 for NNSs).

Speaker 3. The difference between guises was not very pronounced for this speaker; the overall impressions were mixed too. Regardless of the guise, about half of the participants (both NSs and NNSs) saw her as confident and outgoing, possibly rushing a sorority, while others described her as sheltered, naïve/immature and unsure of herself. Only one NS noticed rLIKE in her speech, but many NSs and even one NNS pointed to the fact that she used the word "pop" instead of "soda" which indicated to them that the speaker was Midwestern. Multiple participants also described her as a college student and a family-oriented person but this could be influenced by the content of her speech (the fact that she was missing her mother and sister when in college). Maggie (f, 21, China) compared her to Speaker 1 and claimed that she belonged to the same type, "traditional Americans". Both NSs and NNSs thought that Speaker 2 was slightly younger in her like guise (17.5 vs. 19.3 for NSs, 18.78 vs. 19.07 for NNSs).

Speaker 4. For NSs, this speaker sounded like the oldest of all and slightly younger in her bare guise: 26.5 vs. 28.2. She sounded almost the same age in both her guises to NNSs, but the overall age estimate was very different: 20.25 and 20.58.

The general image that emerged from all profiles was that of an intelligent, determined, mature and possibly nerdy and/or socially awkward person; some participants described her as adventurous as well. Most participants described her as a white American, but several doubted her race (some suggested she was African American or Hispanic) and native speaker status. While her image was essentially the same across both guises, it was stronger and more detailed in the bare guise, although it was in the bare guise when one NNS suggested she was faking intelligence: "Will graduate a 5th year senior and act like its because there was so much more to learn but actually she just failed a class" (b.p., f, 18). rLIKE in her speech attracted minor

attention from the participants (both its presence in the like guise and its absence in the bare guise), although one NS hallucinated rLIKE in the bare guise.

Speaker 5. This speaker was overwhelmingly profiled as lacking confidence, shy/insecure and nervous across both guises; multiple participants mentioned that he stuttered when speaking and "had difficulty choosing words". Four NSs reported rLIKE overuse. Another common theme was him being "artistic" and a "theatre guy". Even though this theme was at least partially based on the content (the fact that he described his path from "the lights guy" to a stage manager in theatre), most of the participants did not simply mention the fact (as was the case with Speaker 3 being a college student and a family-oriented person) but connected it to some other thoughts about the speaker. For example, Al (m, 18) wrote: "He sounds like lower 30s white guy who was bullied in high school. He sounds perfect for the theatre, total theatre kid." Kate (f, 20) gave a different reason: "He sounds like a theater guy (because he's talking about doing theater tech, but also because he sounds like he is acting a little bit)." Speaker 5 was perceived to be younger in the like guise by NSs (19.6 vs. 23.6) but NNSs did not see much difference in age between guises (22.21 in like, 22.7 in bare guise).

Speaker 6. The overall impression of this speaker across-guise, shared by both NSs and NNSs, was that of a friendly and calm girl, relatable, positive and fun to be around, possibly even "rebellious, wild, free-spirited" (Ryan, m, 22, Vietnam). Several NSs noted that her "older" voice did not match her underdeveloped storytelling skills. Curiously, none of the participants noticed rLIKE in her speech, although one NS hallucinated it in the bare guise. To NSs, she sounded much younger in the like guise (20.5 vs. 25); the NNSs agreed (18.58 vs. 19.87).

Concluding the analysis of the freely created profiles, I would like to underscore that different guises did not result in drastically different profiles of any of the speakers. However,

for the first two of the speakers, the like guise was more distinct and pronounced, while for other speakers this was not the case. The reasons for this difference are not immediately evident, although it is possible that it was partly due to the fact that the contents of the speech of the first two speakers were completely devoid of personality, so the listeners were forced to rely on their speech (voice, syntax, word choice, various discourse phenomena and vernacular elements). When profiling speakers 3-6, listeners also had content to partially rely on.

The profiles of the first two speakers differed from the other profiles in another aspect as well. Only these profiles may be undoubtedly described in terms of a distinct social persona ("sorority girl" for Speaker 1 and "frat guy" for Speaker 2), with both terms taken directly from listener descriptions. Notably, even though more NSs mentioned these personae with regard to these speakers, some NNSs did that as well, which indicates their awareness of the existence of such personae. This happened even though the concept of Greek life is specific to American higher education system and does not have direct equivalents in other countries and none of the NNSs in the sample reported belonging to or having tried to join a fraternity or a sorority.

Assigning social personae to the other four speakers involves a certain degree of speculation, but it may nonetheless help with interpreting the results of the second part of the matched-guise experiment. Based on the available data, Speaker 3 may be described as a paler version of the same social persona as Speaker 1, namely a sorority girl. Speaker 4 was generally (although not unanimously) perceived as a nerdy girl. Speaker 5 can be described as an artistic guy (an image almost opposite to the frat guy image of Speaker 2). The profile of Speaker 6 is the most difficult to describe in terms of a social persona because none of the participants proposed a suitable term. However, given the fact that many female participants described her as

relatable and male participants used the words “friendly”, “positive”, etc., I think it is possible to call this image “a girl next door”.

Attribute-Based Ratings

During the second part of the experiment, the participants listened to the same recordings and decided to what extent each speaker could be described as friendly, intelligent, educated, polite, attractive, ambitious, and confident. The judgments were made along a 5-point Likert scale with lower ratings meaning the stronger association between the speaker and a given attribute.

To test whether native and non-native speakers rated the speakers similarly and if any of the other factors (guise, individual speakers, attributes) affected the ratings they gave, I ran a Repeated Measures ANOVA with rating scores as a dependent variable, Group (NSs and NNSs) and Guise (LIKE and bare) as between-subject independent variables (x 2 levels each), and Attribute (x 7 levels) and Speaker (x 6 levels) as within-subject IVs. Examination of the descriptive statistics and of data visualizations showed that the data were positively skewed, and thus not normally distributed, and did not have equal variances. Mauchly’s test of sphericity was statistically significant ($p < 0.05$) for some of the variables, so a correction factor (Greenhouse-Geisser) was used. No important deviations from normality and homogeneity of variances for the residuals were discovered.

The two-way interaction between Attribute and Group was statistically significant ($F_{6,239} = 4.500$, $p = .001$, $\eta^2_p = .083$), although the effect size shows that this interaction accounts for a very small amount of variance. The interaction between Attribute and Speaker was also significant, with a slightly larger effect size ($F_{30,703} = 9.048$, $p < .001$, $\eta^2_p = .153$) The main effect of Attribute was significant and had much stronger effect size ($F_{6,239} = 24.554$, $p < .001$, $\eta^2_p =$

.329). Finally, significant main effect was observed for the Speaker variable ($F_{5,220} = 4.500$, $p < .001$, $\eta^2_p = .140$).

While it is standard practice to report only the results of within-subject effects when performing RM ANOVA, I consider it important that the Group variable was not statistical, while Guise was ($F_{1,50} = 8.021$, $p = .007$, $\eta^2_p = .138$). This indicates that ratings given by the listeners depended more on the presence or absence of rLIKE (independently of other factors, as no interaction between Guise and other variables was statistically significant), on an individual speaker, on the actual attribute in question, and on each of those attributes applied to each individual speaker. I would speculate that, given the exceptionally small effect size, the significant interaction between Attribute and Group may be seen as quantitative support for the qualitatively observed differences in the way NSs and NNSs understand some of the attributes (e.g., friendliness or politeness) themselves (see the case of Chinese participants understanding friendliness as lack of open hostility or merely as slower pace of speech discussed earlier in the chapter).

With regard to most of the attributes, the difference between the ratings given by NSs and NNSs was only quantitative, not qualitative. It can be illustrated by 3-D area plots for perceived educatedness (see Figure 5.2): while NNSs were generally "kinder" towards the speakers and tended to judge them as more educated than their NS peers thought they were (it is reflected in the smaller size of the object representing NNS-given ratings), the slopes (i.e., the difference between ratings given in the like and bare guises) for most of the speakers look similar or even identical. Of course, the level of overall similarity between NS and NNS judgment varied between different attributes, but it was due to the intertwined effects of attribute, guise and speaker.

In order to further analyze the effects that guise, individual speakers and attributes in question had on the judgments made by the participants (and reflected in rating scores), I looked at whether the ratings given to speakers were lower (i.e., positive, as the score of one described an attribute as matching a given speaker perfectly) or higher with rLIKE present. The results can be seen in Table 5.4.

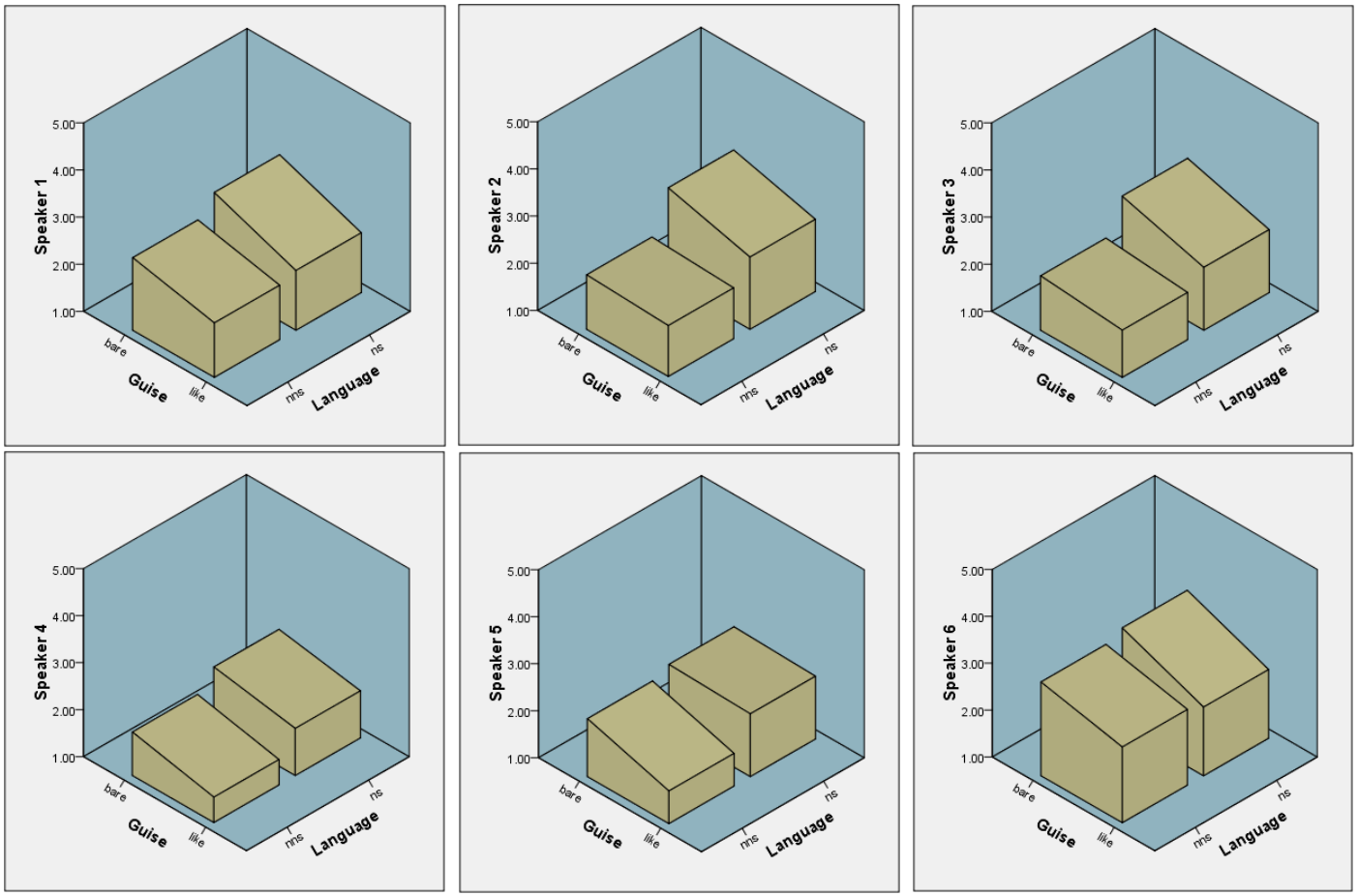

Figure 5.2

3-D Area Plots for Perceived Educatedness of the Speakers

Table 5.4
The Effect of rLIKE Presence on Perceptions of the Speaker by NSs and NNSs

	S1		S2		S3		S4		S5		S6	
	NS	NNS	NS	NNS	NS	NNS	NS	NNS	NS	NNS	NS	NNS
Friendly	√	√	x	x	√	√	√	x	x	√	x	=
Intelligent	√	√	x	x	√	x	x	x	x	√	x	x
Educated	√	√	x	x	√	√	x	x	√	√	x	x
Polite	√	√	√	x	√	√	x	x	√	√	√	x
Attractive	√	√	√	x	√	√	x	x	x	√	√	x
Ambitious	x	x	√	√	√	√	√	x	x	x	x	x
Confident	√	x	x	x	√	√	x	x	x	x	x	x

Note. The √ symbol represents positive effect (i.e., the speaker is perceived as more friendly, intelligent, etc. in the like guise), x represents negative effect, = means that the speaker received the same rating regardless of the guise. The areas of agreement between NSs and NNSs are marked with gray color.

As evident from Table 5.4, while the overall level of agreement between NSs and NNS (marked with gray color in the table) is high, it is not consistently high across all speakers or across all attributes. The only attribute which elicits unanimous agreement from both NSs and NNSs is educatedness; confidence and ambitiousness are rated similarly for five out of six speakers. The friendliness was the attribute which was similarly perceived only with regard to the first three speakers. However, the most important finding here was that the presence of rLIKE appeared to affect the perception of different speakers differently. In other words, there was no attribute for which the ratings were affected only negatively or only positively across all speakers. For example, the presence of rLIKE made Speakers 1, 3, and 5 sound more educated but produced the opposite effect on Speakers 2, 4 and 6, and this effect was the same according to both NSs and NNSs.

Furthermore, it appeared that not only were certain attributes seen as more or less associated with a given speaker depending on the presence of rLIKE in his or her speech, but some speakers (1 and 3) were likely to be judged more positively across the board when using rLIKE, some more negatively (4 and 6), while some produced mixed impressions, e.g., “gaining”

in some qualities and "losing" in others (speakers 2 and 5). While a new experiment would be necessary to confirm it, I hypothesize that these differences in listener perceptions of the speakers are rooted in the social personae associated with each of the speaker. Similarly to the way the participants of D'Onofrio's (2015) experiment expected to hear TRAP-backing when they thought they were listening to a "valley girl", the listeners in my study would be likely to expect to hear rLIKE from a speaker they perceived to be a young white "sorority girl" and unlikely to expect it when listening to a "nerdy girl". It is notable that both speakers who were described as belonging to the sorority girl type of person (1 and 3) "gained" in four of the qualities (friendliness, educatedness, politeness, attractiveness) according to both NSs and NNSs and in intelligence and confidence according to NSs (all despite the fact that the listeners did notice rLIKE in their speech). It seems possible that this effect could occur due to the fact that in their like guises these speakers matched the listener expectations perfectly while in the bare guise they "did not sound right" which would lead to lower ratings on most of the attributes. Similarly, stereotypical nerdy girls and (to a lesser extent) girls next door are not expected to overuse rLIKE, which could explain why speakers 2 and 6 had lower ratings on most of the attributes in their like guises.

The speakers whose ratings were mixed (i.e., rLIKE presence had a positive or negative effect on ratings depending on the attribute), 2 and 5, were both male but represented very different types of social personae ("frat guy" and "artistic guy"). It can be speculated that in their case, the expectations based on gender prevailed: after all, the only male social persona associated with rLIKE usage and mentioned during the attitudinal interviews and questionnaires (and mentioned only once) was that of a gay man, and none of the two male speakers in this experiment was profiled as gay. It appears that female social personae may be divided into those

expected to use rLIKE (e.g., “sorority girl”) and those who are not (e.g., “nerdy girl”), while no such division can be observed for male social personae. It supports the hypothesis I proposed earlier, based on the composite profiles of the speakers, that the perceptions of male and female speakers may differ with regard to rLIKE usage.

Conclusion

Overall, the results of the experiment only partially supported the findings of previous studies. The study that reported “the knee-jerk ‘like is bad’ reaction” (Hesson & Shellgren, 2015) used a single female speaker, and two of the studies that found the positive effect of rLIKE on solidarity-based traits and negative on the status-based traits used either a single female speaker as well (Maddeaux & Dinkin, 2017) or written stimuli (Buchstaller, 2006). The only previous study that used different speakers (in terms of age, gender, etc.) was Dailey-O’Cain's (2000), but the reported results also focused on the different effect rLIKE had on solidarity-based and status-based attributes. The results of my experiment showed that the attitudes towards rLIKE may present a more complex picture, with listener judgments fine-tuned to differ depending not only on the presence of rLIKE and not only on the type of the attribute in question (solidarity or status-based) but the perceived gender and, even more importantly, the social persona of a given speaker.

The most important finding, however, is that NNSs seemed to be following the same intricate pattern of judgment as NSs. Even though the actual ratings given by NSs and NNSs may differ numerically, the agreement on the type of the effect (positive or negative) rLIKE has on the perception of a given attribute in a given speaker is overwhelmingly high: NSs and NNSs agree upon over 70% (30 out of 42) “cases” (Attribute x Speaker) to judge. Furthermore, the pattern of agreement seems to be systematic rather than random, as the highest levels of

agreement between NSs and NNSs (6 out of 7 cases) appear in the judgments of speakers 1 and 3 which both, according to the free profiling trial, represent the same social persona, a sorority girl.

CHAPTER SIX

CONCLUSION

This dissertation explored the ways in which non-native-speaking international students on an American college campus use remarkable LIKE, understand it, assess its syntactic placement in the speech of others, think about it and about people who use it. I used a combination of \ experimental techniques (syntactic judgment and matched-guise experiments) and other methods (interviews and methods) and employed qualitative and quantitative approaches to data analysis. While the primary focus of my study was on the twenty-six non-native speakers of English, native speakers were recruited to participate in some of the tasks so that a native speaker reference point (of syntactic judgments or with regard to attitudes) could be established for the subsequent analysis. To analyze patterns of syntactic placement of rLIKE, I also used data from the IHELP-MI corpus, which contains the speech of young people from Michigan State University: the same institution that my own study participants were attending at the time of data collection.

In this final chapter, I will first summarize the main findings of this study, then discuss its implications for second language pedagogy and ways of spreading discourse-pragmatic awareness among English learners. Then I will draw the reader's attention to the limitations of this study and conclude with suggested avenues for future research that this study has opened.

Main Findings

The most clear and undisputable finding of this study is that undergraduate international students who have already spent some time on an American university campus use remarkable LIKE in their own speech. At the same time, the frequencies with which they use it vary

drastically, from two to almost eighty tokens per 1,000 words, with the average frequency being around twenty-seven.

NNSs in my study were also acutely aware of rLIKE as a discourse-pragmatic phenomenon. Most of them said that they had not heard it before arriving to the United States, which may be the case, especially if a non-native speaking instructor without much exposure to native speech was their only source of L2 input, or which might mean that they heard rLIKE so infrequently that had not noticed it in the input. The important fact is that they all noticed it in the speech of their native-speaking classmates and other people around them and were able to interpret it as a discourse feature, namely. as something that does not have lexical meaning, rather than as any of the unremarkable functions of LIKE (verb, comparative complementizer) that they had known before.

Despite the high level of inter-individual variation, as a group, NNSs demonstrated remarkable native-likeness in usage patterns and judgments of syntactic placement of rLIKE. As far as usage is concerned, the distribution of all tokens across functions in NNS data and in native data in IHELP-MI was strikingly similar: for example, tokens of clause-initial discourse marker LIKE comprised 22.7% of all native rLIKE tokens, while among NNSs the proportion was 25.3%; with regard to the clause-medial discourse particle, it was 38.4% and 44.8%, respectively. The same type of similarity was observed for the distribution of clause-medial tokens across syntactic positions in native and non-native data.

This supports Schleef's (2013) conclusion based on extensive review of variationist studies with non-native speakers of various languages, that “the patterns of variation used in the interlanguage of L2 learners do approximate those of native speakers” (p. 306). Crucially, my study permits a cautious extension of this claim to perception-based judgment patterns as well.

The order of naturalness created by NNSs with regard to different syntactic positions of rLIKE generally matched the order created by NSs and, which, in its turn, was consistent with the order that could be expected judging by the diachronic and distributional hierarchy of rLIKE functions and syntactic positions demonstrated by D'Arcy (2017). While native and non-native orders of naturalness were not exactly the same and differed with regard to syntactic positions within the verbal domain (e.g., between or within *to*-infinitive, etc.), certain critical differences that were expected were, indeed, observed in both native and non-native judgments. For example, both groups judged sentences with LIKE in clause-final position, which is not present in American English, to be very low in naturalness. Another example is the difference between judgments of clause-medial LIKE in the DP context (highly natural in both groups) and in the nP context (significantly lower), which is consistent with D'Arcy's corpus data, where LIKE appears in the well-established DP context with much higher frequency than in the newer nP context.

Furthermore, NNSs seemed to be following the native pattern with regard to a different type of judgments, namely, the social judgments about rLIKE users that they demonstrated during the matched-guise experiment. I found that NSs and NNSs tended to agree on the type of effect (positive vs. negative) the presence of rLIKE in speech had on the judgment of how well a certain attribute (e.g., friendliness or intelligence) described a certain speaker. Importantly, this effect seemed to depend not just on the presence of rLIKE but on speaker gender and social persona as perceived by the listeners. For example, a speaker perceived as a sorority girl was more likely to be described as friendlier and more intelligent when she used rLIKE, while someone perceived as a nerd could be described as less friendly and intelligent in the LIKE-guise. Given the intricacy of this pattern, the agreement between native and non-native speaker judgments is best described as systematic.

The analysis of variation among the NNSs did not result in any statistically robust findings with regard to social factors that could have affected rLIKE usage and the nativeness of syntactic judgment patterns. However, length of residence emerged as a possibly important factor. Even though longer LOR could not serve as a reliable predictor of higher frequency of rLIKE in NNS's speech (some speakers with longer LOR were among the least frequent rLIKE users), almost all of those who demonstrated the most native-like patterns of syntactic judgment also reported above-average LOR (i.e., over 40 months). Interestingly, beliefs appeared to be a potentially important factor affecting rLIKE usage. All of the top-six speakers who used rLIKE at an above-average rate thought that rLIKE was a specifically American phenomenon, "an American thing". At the same time, those who did not recognize it as such, were less likely to use rLIKE frequently. This finding adds to the mounting evidence that NNSs may use socially salient vernacular features (at various level of language, including phonology and grammar) as a way to signal their belonging to the native-speaking culture (Sharma, forthcoming).

The analysis of speaker beliefs about rLIKE specifically resulted in interesting insights as well. Native speaker beliefs expressed in this study were consistent with previously reported findings (Fox Tree, 2007) as well as with "myths about LIKE" summarized by D'Arcy (2017). Thus, NSs recognized rLIKE as a distinct discourse phenomenon, associated with young white people, especially women, as a symbol of inarticulateness and as a word without any meaning in particular that should be avoided in any kind of professional and/or formal situation. Most NNSs, including those who did not frequently use rLIKE themselves, however, demonstrated a surprisingly detailed understanding of rLIKE as a discourse element and described it as having multiple functions, from marking hesitation to linking parts of the sentence, and discussed how it can be used to describe things or create an informal atmosphere. Many of the NNSs, as I

mentioned earlier, also saw rLIKE as a specifically American phenomenon, which may have prevented them from observing the differences in rLIKE frequency in formal and informal context. Only a few NNSs said that it should be avoided in professional or other similar settings, and out of those, two reported being told this explicitly by public speech coaches. Finally, many NNSs, none of whom was majoring in any language- or linguistics-related area, reported being intentionally observant about the ways in which native speakers around them used rLIKE and other vernacular elements and trying to imitate native speaker behavior in that respect.

Pedagogical Implications

Remarkable LIKE, as well as many other vernacular features, is not typically taught in the classroom. This is a common knowledge assumption among English teachers and second language researchers worldwide (Algouzi (2015) was able to confirm this assumption at least with regard to textbooks and materials used in Saudi Arabia). Given the syntactic and semantic (though not always pragmatic) optionality of it as a discourse marker or particle and existing synonyms for it as an approximator (e.g., *about*) and a quotative (e.g., *say*, *think*), which are usually learned early, its absence from the classroom materials is understandable.

However, the high number of NNSs who reported never hearing rLIKE before arrival to the U.S. suggests that there is a potential need for more authentic input that would introduce learners to rLIKE and possibly other highly frequent elements of the common vernacular (i.e., vernacular English features that are not restricted to any specific group of speakers). It could be even more beneficial if this input provided learners with opportunities to notice the stylistic differences in rLIKE usage and observe that it is less frequently used in formal contexts. The lack of stylistic awareness with regard to rLIKE characterized many prolific and native-like rLIKE users in my sample, which suggests that, while the acquisition of rLIKE usage patterns is

achievable in the immersion context with plentiful input, stylistic awareness potentially requires some guidance. As Meyerhoff and Schleef (2012) found a similar tendency with regard to another socially salient variable, (ing), in Scottish English (style was a significant constraint for native-speaking but not for Polish-born teenagers), it is likely that the results I obtained are not specific to rLIKE but point to a larger trend. This underscores the importance of pedagogical intervention aimed at promoting stylistic awareness with regard to the stigmatized variants of socially salient variables among NNSs.

Because rLIKE is overwhelmingly perceived as an American phenomenon, NNSs may, in fact, overuse it in formal situations in order to "blend in" and demonstrate their American identity. Given the negative beliefs about rLIKE many native speakers possess and the fact that its usage may even affect the likelihood of being hired (Russell et al., 2008), stylistic awareness about rLIKE becomes a truly important issue that might need to be addressed even in the classroom.

Finally, the findings of the matched-guise experiment, that is, the fact that the presence of rLIKE in the speech of one type of social persona had a drastically different effect on listener attitudes than the presence of rLIKE in the speech of another type of social persona, have implications for using matched-guise technique in sociolinguistic research in general. Studies that employ this technique often use only one speaker, which may result in generic conclusions, such as "LIKE positively affects solidarity-based evaluations and negatively affects status-based evaluations", which would obscure the fact that for some types of social personae, LIKE may have the exact opposite effect.

Limitations

While all studies in the area of discourse-pragmatics have their limitations, a study that ventures into the previously unresearched or under-researched terrain has many, and a transdisciplinary study might have even more.

First of all, there are limitations typical for many second language studies: relatively small sample size (it especially applies to the experimental tasks) and self-selection of the participants (i.e., they choose to respond to a recruitment call). While small sample size limits the reliability of statistical tests, self-selection implies a certain level of linguistic security on the part of the participants. In other words, they are not afraid to be interviewed and perform tasks in English, and as such, the findings based on their speech data may not necessarily be applicable to less linguistically secure international students on the same campus. It is also evident that findings obtained with regard to student data cannot be generalized to all other non-native-speaking populations, such as immigrants or even graduate students at the same university, due to potential differences in not only prior education but also the type of input they receive on a regular basis.

There are also limitations imposed by the study design and the difference in procedure for native and non-native speakers. While the fact that NSs were surveyed and tested online and I met with the NNSs in person did not seem to have an apparent effect on the patterns discovered based on experimental data, there is no doubt that surveying NSs instead of interviewing them prevented me from asking follow-up questions and pursuing certain themes in-depth.

Because the data were collected at one point in time and this imposed a LOR threshold on the participants, no conclusions can be made about the acquisition of rLIKE as a longitudinal process and the development of usage and judgment patterns.

Finally, it would have been interesting to compare the behavior of speakers of different L1s with regard to rLIKE based on the presence or absence of rLIKE equivalents in those L1s. However, it could not have been done within this study because of the lack of previous research on the matter, so the role of L1 in NNSs's behavior remained unexplored.

Future Research

The possibilities for future research are theoretically innumerable, but I would like to focus on two major avenues for future research that this dissertation opens.

First of all, as this study was the first of its kind, a conceptual replication, and possibly partial replication of the experimental parts, would help confirm and/or clarify the findings of the study. It would be especially interesting to see a similar study conducted with a different population of NNSs, such as immigrants, highly proficient speakers in a non-immersed setting, etc.

A new line of research could involve conducting a study focused on the process of acquisition of rLIKE and attitudes towards it. Investigating the acquisition of rLIKE would require interviewing and testing participants several times, for example, during the month of arrival, after the first year on campus, after the second year, etc. Investigating the acquisition of attitudes could be based on longer, more in-depth interviews with selected speakers (i.e., those who already expressed native-like negative attitudes and/or stylistic awareness) focused specifically on what sources and people could have contributed to the acquisition of those attitudes.

For now, however, this dissertation has made what I hope is a valuable contribution to the literature on discourse-pragmatic variation in non-native speech, strengthening the

transdisciplinary connections between second language acquisition and language variation and change research.

APPENDICES

APPENDIX A

Background Survey for Non-Native Speakers

Basic information

Age (in full years): _____
Gender (circle/fill in):
Male Female Prefer not to say Other _____
Current year at MSU (circle/fill in):
Freshman Sophomore Junior Senior _____
Intended major(s): _____
Intended minor(s): _____
Have you ever taken a course in sociolinguistics?
Yes No Not Sure
Have you ever taken any courses in linguistics or language learning? (list either as course codes or descriptions): _____

Geography

1. Country of birth: _____
Did you grow up there?
Yes No (please indicate where: _____)
2. Arrival in the United States (month and year: e.g., August 2017): _____
3. After you moved to the United States, have you only lived in Michigan?
Yes No (please specify state, city and length of stay_____)
4. Have you ever lived in an English-speaking country for more than a month before coming to the U.S.?
No Yes (please specify country and length of stay _____)

Linguistic background

1. What language(s) did you speak at home growing up? _____
2. What language(s) did you speak at school in your home country? _____
3. When did you start learning English? (circle/fill in)
Before Elementary School In Secondary/High School Other _____
In Elementary School At College/University
4. If you remember, please specify how old you were when you began learning English. (in full years) _____
5. Have you ever been to an English immersion school/camp in your home country?
Yes No
6. Have you ever studied abroad in English or for the purpose of learning English?
No Yes (please specify: _____)
7. Were you required to take classes at the English Language Center when you arrived at MSU?
Yes, I had to spend a semester or more studying just English
No, I started taking regular classes right away
I was allowed to take regular classes, but also required to take some English classes
8. What was your most recent TOEFL score?
_____

Did not take TOEFL, but took another standardized test (please specify test and score: _____)

Did not take TOEFL or another standardized test

Prefer not to say

9. Do you speak any language(s) other than your native language(s) and English? If yes, please list the languages you speak, starting with the one in which you are most proficient.

_____

_____

_____

APPENDIX B

Background Survey for Native Speakers

1. Please choose a pseudonym (nickname) we can use throughout the study to identify your answers. Choose a name that does not resemble your real one in any way as this is crucial for the protection of your identity. Do not choose nicknames you routinely use online or otherwise. (Text box)
2. How old are you? (Selection)
3. What is your gender? (Select one: Male, Female, PREFER NOT TO SAY, Other)
    a. If other, please specify (Text box)
4. How do you identify racially? (Select one: Asian/Pacific Islander, Black or African American, Hispanic or Latino, Native American or American Indian, White/Caucasian, Other, Prefer not to say)
    a. If other, please specify (Text box)
5. What is your current year at MSU?
    a. If other, please specify (Text box)
6. What is/are your intended major(s)? (Text box)
7. What is/are your intended minor(s)? (Text box)
8. Have you ever taken a course in sociolinguistics? (Select one: Yes, No, Other)
    a. If other, please specify (Text Box)
9. Have you ever taken any courses in linguistics or language learning? Please list those (you may use codes or course names/descriptions).
10. Were you born in the United States? (Select one: Yes, No)
    a. If Yes
        i. Question #11
    b. If No:
        i. In what country were you born? (Text box)
11. Were you born in Michigan? (Select one: Yes, No)
    a. If Yes
        i. Did you only live in Michigan from birth until the age of eight? (Select one: Yes, No)
            1. If Yes
                a. Question 12
            2. If No
                a. In what other state did you live until the age of eight? (Text box)
        ii. Have you ever spent a consecutive period of a year or more outside of Michigan? (Select one: Yes, No)
            1. If Yes
                a. Question 12
            2. If No

a. Question 13

b. If No

i. In what state were you born? (Text Box)

1. If not blank, did you live in your home state from birth until the age of eight? (Select one: Yes, No)

a. If Yes

i. Question 13

b. If No

i. Question 11(a)(i)(2)(a)

12. In what other state have you spent a year or more? (Text box)
13. Are you a native speaker of English? (Select one: Yes, No)
    a. If Yes
        i. Question 14
    b. If No
        i. What is your native language? (Text box)
14. Are your parents native speakers of English? (Select one: Yes, both of them, No, neither of them, Only my mother, Only my father, I am not sure)
    a. If both of them
        i. Question 15
    b. If only Mother
        i. What language(s) does your father speak natively? (Text box)
    c. If only Father
        i. What language(s) does your mother speak natively? (Text box)
    d. If I am not sure
        i. Question 15
15. Are all your grandparents native speakers of English? (Select one: Yes, No, none of them is, I am not sure, Some of them are)
    a. If some of them are is not selected
        i. Question 16
    b. If some of them are is selected
        i. Please specify what language(s) do/did your grandparents speak natively? (Text box)
16. Do you speak any languages other than English? (Select one: Yes, No, Yes, but just a little)
    a. If No
        i. Question 17
    b. If Yes/Yes, but just a little
        i. Please list the languages you speak (other than English), starting with the one in which you are most proficient. (Text box)
        ii. How would you describe your proficiency in the first language you listed in the answer to the previous question? (Select one: Beginner (just began to take a class or know a few common phrases), Novice (can maintain basic conversation, understand signs and short simple texts), Intermediate

(can comfortably communicate with native speakers on non-specific topics and/or read newspapers and light fiction), Advanced (can speak fluently, participate in communication on complex topics, read almost any type of text), Near-native (cannot be easily distinguished from native speakers, except maybe by slight accent), Bilingual (grew up speaking this language along with English, not distinguishable from native speakers in any way))

iii. What language(s) do you speak at home? (Select one: English only, English and another language, Another language only, Other)
    1. If English only/Another language only
        a. End of block
    2. If English and another language/Other
        a. Please specify (Text box)

APPENDIX C

Interview Protocol for Non-Native Speakers

Note. Before the interview, participants will be asked if they agree to be recorded.

1. Can you tell me a little bit about yourself? Where were you born, where did you grow up?
    - Tell more about the place
    - Tell more about the family
2. How did you end up at MSU?
    - Education prior to MSU: only school or some college?
    - How was the decision made?
    - Why the US/Michigan?/MSU?
    - Did you have to go through the ELC?
3. How do you like it here at MSU?
    - Do you live on campus?
    - Roommates?
    - Friends: how many? NSs/NNSs?
    - Extracurricular activities
    - Classes, professors
    - Communities based on native country/language
4. What are your plans for the future?
    - Go back home, stay in the US, move to a third country?

Note. A specific attempt must be made to bring up an emotion-provoking and, hence, vernacular-friendly topic. These will vary between speakers and may include childhood hobbies, first day on campus, pets, sports, and other topics that a speaker seems to be interested in. Also, special attention should be paid to follow any identity-revealing topics the speakers bring up.

APPENDIX D

Instructions and Stimuli for the Syntactic Judgment Task

Instructions

The following instruction was given to the participants (both on screen and aurally): “You will listen to a number of sentences. For each one, please indicate how natural the sentence sounds to you on a scale from 1 (unnatural) to 5 (perfectly natural). In other words, decide whether the sentence sounds like something that you might hear in everyday casual speech, for example, on campus. It’s fine if you do not understand the broader context of a given sentence.”

Stimuli

1. LIKE we spent our childhoods trying to be different.
2. I don't see my dad very often LIKE he knows what's going on.
3. So, um, my mom, she served on a, um an agriculture LIKE advisory board.
4. My goal isn't to make money… My goal is to LIKE make a difference in people's lives.
5. I kind of wanted to LIKE prove to other people that you know, I could take on the challenge.
6. People will just come up and LIKE sit with you and get to know you and stuff.
7. I'm not so much into LIKE European History and that type of stuff.
8. I visit her like four or five time before LIKE to ask her questions about applications and all that stuff.
9. I've always been up here for sporting events, and LIKE to visit people that have been up here.
10. I mapped out LIKE where I have to go for my classes yesterday.
11. We all live on the same LIKE street.
12. And then you go through dungeons and LIKE fight monsters.
13. My sister's LIKE twenty-seven.
14. And um, the fraternity guys are getting in groups of ten and doing LIKE challenges.
15. Halloween is separate, but for LIKE, the week to week stuff we have themes.
16. LIKE last week we dressed up as superheroes* because of our philanthropy and stuff like that.
17. That depends on LIKE which school they wanted to go to.
18. But I also like that I can LIKE stay here for school.
19. But he doesn't do his laundry for LIKE two months at a time.
20. If you, if it gets bigger, then you could LIKE be with them from the start and I don't know make their image I guess.
21. I like the big city vibe, LIKE the whole museums and stuff.
22. It was considered on campus housing, and LIKE everything was through one agency.
23. One of my cats meows so much ‘cause LIKE he’s really picky and everything.
24. Whoever has the most after LIKE all the people go is the winner.
25. The rafting guy actually made me sit up the front LIKE, ‘cause I didn’t have a paddle.
26. We need to smarten it up a bit LIKE.
27. He LIKE was so happy to take a bath.

28. I had a crush on LIKE him.
29. I've caught trout that LIKE are small.
30. She's very aware of her feelings but is un-like-sympathetic to others.
31. We all are LIKE down to earth type of people.
32. But I've never actually LIKE been up here.

Fillers

33. She pretty much stays in her room.
34. Homecoming, that was always fun.
35. We don't really step on each other's toes too much, so.
36. We just you know, hiked on mountains and the ocean and stuff.
37. I'm doing a lot less work than I was senior year.
38. My one friend is a meat man as we call him at Meijer.
39. I actually met him at a football game.
40. But I have work, so I don't think I'm gonna make it.
41. I tried volleyball for a little bit.
42. I'm living with four other college aged girls in a house together.
43. Well we went to the same schools since kindergarten.
44. And they have rock stations too, but depends what you listen to.
45. I know that he got kicked out of a Catholic school
46. I mean I kept some of my friends from high school.
47. I guess we just didn't have the money for it or whatever.
48. I don't really know my dad's side of the family.
49. I think that's the only time I ever got really mad.
50. I mean people would never actually use it.
51. Well I've always been into history
52. She got her two different masters at MSU.
53. And I was hoping she'd go to State cause she's really awesome.
54. It's not too bad of a walk, though.
55. Well, actually I lived back down in Virginia also.
56. I know a couple people, I don't really communicate at all with them though.
57. I was friends with everybody, so.
58. I made a lot of friends with some people that were younger than me.
59. I met a lot of friends there that I'm really good friends with.
60. I take the bus on Fridays because my friend takes the bus back.
61. We couldn't really have sick days or personal days or anything.
62. The last day that you could drop classes I dropped it.
63. My mom went to Michigan State, and that's where her and my dad met.
64. I know my dad lived in Indiana for a while.

"Ungrammatical" items

65. They went yesterday to the movies.
66. My friends drink always coffee from Starbucks.
67. She asked me what is the time.
68. My roommate explained me the math problem.
69. The beer light is the nastiest.

70. This pizzas don't taste good.
71. Class start at eight usually.
72. I don’t know where is the restroom.
73. If it will be sunny, we’ll go to the park.
74. I don’t like the musics here.
75. My parents thinks I’m no good at sports.
76. My sister tomorrow is going to a concert.
77. I no dance, I just sing when I drink.
78. Our friends goes to Paris tomorrow.
79. My roommate buy groceries every Saturday.
80. My car no has gas to get home.
81. She must to go to class tomorrow.
82. Two Coke and ice, please.
83. He have two cars and a bike.
84. I drink coffee with three sugar only.
85. Here is the number, call back to him.
86. Please give me a pen, the one blue.

Stigmatized vernacular

87. I don’t know nothing about that.
88. You know, me mother never gives a damn.
89. They says it’s not a good idea.
90. These white shirts are dirty - they need washed.
91. She ain’t good enough for you!
92. I seen him cross the road and almost get hit by a truck.
93. We might could think about going to the concert.

Manipulated idiomatic expressions

94. My grandparents still have a white and black TV.
95. I don’t like glue, tape is better, it’s sort of dirty and quick.
96. The waiter brought out my cheese’n’mac.

APPENDIX E

Materials for Matched-Guise Profiling Task

Free profiling task instructions

You will hear six stories told by different people. After each story, please try to guess the speaker's age and write everything you can say about the person who told the story. Bullet points, words divided by a comma, or complete sentences are all fine: choose the style that is most comfortable for you. There is no "right" description your answer will be compared against; it is only your personal impression that is valuable. You do not need to justify your impression in any way, but you certainly can if you want to. Take as much time as you need. When you are finished with a story, please press the button and listen to the next one.

Attitudinal task instructions

You will hear six stories told by different people. Do not worry if you recognize the stories. After each story, please use the 5-point scale to indicate the degree to which certain qualities may be associated with the speaker.

Stories (LIKE guise)

1. A Cap and a Hat (female speaker; 7 tokens of rLIKE)

Um, I feel like a cap is a -- more specific. A cap is a type of hat. LIKE a hat could be all these different sorts of things that you put on your head. LIKE it wouldn't be, LIKE a blanket that you put on your head but LIKE different things that you can put on your head to cover your head. And a cap would be, I would think more LIKE a baseball cap. Or, um, LIKE one of those grandpa's caps. Yeah. Because a hat is much broader. So you could use different things, and caps, usually LIKE baseball caps are probably made out of the same thing.

2. A Townhouse and a Duplex (male speaker; 4 tokens of rLIKE)

Townhouses are LIKE a hu- +huge --bigger group, cause I lived in a couple townhouses. And it was always LIKE a group of a couple of them together. And duplexes are just two, usually. From what I can tell. Yeah. I've seen with garages, also without garages, but duplexes, they're just houses, so they do have garages also. Um, townhouses they're also usually more s- -- located in LIKE apartment complexes. I've noticed that. I k- +kind -- kind of consider them LIKE apartments, in that group.

3. Parting with the Family (female speaker; 8 tokens of rLIKE)

Me and my sister are really close, and so I'm probably gonna go home LIKE every other weekend probably to see my sister, so. Yeah, she wrote me a big ol' letter before she left and

everything. My mom-- she's very LIKE an emotional person. So, LIKE we -- she started crying LIKE a week before I even left or while we were LIKE going dorm shopping and stuff, so. I WAS LIKE, "Mom, jeez." So I've got everything. I s- +saw -- I thought I was missing it -- something, but I got everything. I was pretty proud of myself for remembering everything. But all that's in my fridge is LIKE some pop, a water, and LIKE a packet of ranch or something like that. 'Cause I ordered a pizza the other day so.

4. Future Career (female speaker; 5 tokens of rLIKE)
I think I only applied to M S U in LIKE August. And I got in right away so I said, I don't need to even apply anywhere else 'cause that's where I want to go. Anthropology. Because I love people not necessarily the people around me so much. But just all the different kinds of culture worldwide and how they affect how people live. Mm I want to be LIKE kind-of a humanitarian LIKE person that I'm not sure I d- -- I like Africa a lot so maybe go to Africa and teach I think. Um i- +it -- it depends if -- if I'm LIKE in a different country or something. I don't know LIKE speak English as a second language or something. Um but here If I stayed here, I'd probably be a professor or something of anthropology.

5. Stage Lighting (male speaker; 7 tokens of rLIKE)
Well my brother is an actor. And he likes (gram) to do the theatre. So um, he told me that Mrs. James*, the director needed someone to do lights my freshman year. And I WAS LIKE, "Okay, I'll go see what it's about." It was totally random. And I just thought -- I liked doing lights. I did two shows. 'Cause there's a musical, and a play every year. So I did lights for both of them. And then, um the next year she just came up to me. She's LIKE, "We need a stage manager. The other one left. So I WAS LIKE, "Okay." I di- +didn't -- So I had no idea what I was doing. LIKE I didn't even know what it was I was supposed to do. Because I was a part of the last productions but I – LIKE I was just the light guy. I didn't know very many people uh yeah. And, LIKE 'cause you're supposed to uh at first I didn't know LIKE I'm supposed to direct other people to do stuff. So I was trying to do everything myself. And it took a long time to learn. But by my senior year I was -- I was teaching bunches of people how to do it.

6. Haircut (female speaker; 4 tokens of rLIKE)
I always cut my sister's hair. There was, one time she had it LIKE wrapped around a comb, so I had to LIKE, tug it, and it wasn't coming out so I cut it so she had to -- we had to ask Mrs. Delany* if she could wear a bandana because it was just little hairs poking up. And then there was another time where my sister and I were playing hair salon and I cut her hair, LIKE one side, just one side. And we flushed it down the toilet, and my dad brought us home, and he hadn't noticed. And my mom WAS LIKE, "What did you do?"

APPENDIX F

Protocol for Attitudinal Interview with Non-Native Speakers

Note. Before the interview, participants will be asked if they agree to be recorded.

1. Here is the profile you created for one of the speakers in the first listening task. Do you remember what the story was about? Let's listen to it again. Now can you try to remember why you think this person is young/educated/intelligent, etc.?
2. (The analysis of the second profile)
3. Do you have an idea about the purpose of the study? What do you think we are looking at?
4. We are researching how people use the word LIKE in English other than as a verb. For example, when somebody says: "And I was like 'Oh my God!', or 'And I hate like math'". We also want to know how this – let's say non-textbook - usage may be perceived by other people.
    - Do you think you use LIKE a lot?
    - Do you think your classmates do?
    - Is it easy for you to understand different meanings of LIKE? If yes, has it always been?
    - Do you use LIKE all the time or only in certain situations?
    - Are there situations in which you are deliberately trying not to use LIKE? Why?
    - Have you ever heard anybody discussing the usage of LIKE? Maybe a teacher, a classmate? What did they say?
    - Have you ever read/watched opinions on LIKE in mass media? What were those?
    - Do you think you use LIKE differently than native speakers do? In what way? What kind of people are most likely to use LIKE?

APPENDIX G

Attitudinal Survey for Native Speakers

1. Have you been able to figure out the purpose of the study? What do you think we are looking at? Please write your idea in the box below. (Text box)
2. Debriefing text appears on screen
   We are researching how people use the word LIKE in English other than as a verb. Look at the examples:

   - And I <u>was like</u> 'Oh my God!'
   - I hate <u>like</u> math and stuff.
   - <u>Like</u>, I am always trying to do everything on time.

   Does this ring a bell? We also want to know how this – let's say non-textbook, or vernacular, – usage may be perceived by other people. Please answer a few questions to help us.
3. How often do you think you use "like" in your speech? (Select one: Always, Frequently, Sometimes, Hardly ever, Never)
4. How often do you think other people of your age use "like"? (Select one: Always, Frequently, Sometimes, Hardly ever, Never)
5. Has anyone ever commented on the way you use "like", or the frequency of such usage? Please tell us. (Text box)
6. Have you ever heard anybody discussing the usage of "like" in general? Maybe a teacher, a classmate? What did they say? (Text box)
7. Have you ever read/watched opinions on "like" in mass media? What were those? (Text box)
8. Is there any type or group of people that, in your opinion, uses "like" more than others? (Text box)
9. Are there situations in which you are deliberately trying not to use "like"? Please describe briefly. (Text box)
10. Please rank the following situations from the ones in which you would use "like" freely and frequently (1) to those in which you would try to avoid using like (8). Start dragging the lines and you will see the numbers to appear. (Drag and drop to order 1-8: Talking to friends, Talking to your parents, Talking to your grandparents (or other older relatives), At work, When speaking in class, When talking to a professor one-on-one, Giving an interview to a newspaper, At a job interview)
11. Please let us know what your current job is. Write just the type of job (e.g., "waiter", "cashier", "nanny", "research assistant"), do not mention a particular company. Write n/a, if not applicable. (Text box)
12. Is there anything else you would like to tell us about yourself, or maybe about our experiment? Please feel free to do so here. (Text box)

REFERENCES

REFERENCES

Adamson, H. D., & Regan, V. M. (1991). The acquisition of community speech norms by Asian immigrants learning English as a second language: A Preliminary Study. *Studies in Second Language Acquisition*, *13*(1), 1–22.

Ajmer, K. (2002). *English discourse particles: Evidence from a corpus*. Philadelphia: John Benjamins Publishing Company.

Algouzi, S. (2015). *Discourse markers in Saudi English and British English: A comparative investigation of the use of English discourse markers* (Doctoral dissertation). Retrieved from http://search.proquest.com/docview/1947636036/A41BD2D9390A4A6EPQ/9

Amador-Moreno, C. P. (2012). A corpus-based approach to contemporary Irish writing: Ross O'Carroll-Kelly's use of like as a discourse marker. *International Journal of English Studies; Murcia*, *12*(2), 19–38.

Andersen, G. (1998). The pragmatic marker "like" from a relevance-theoretic perspective. In A.H. Jucker, & Y. Ziv (Eds.), *Discourse markers: Descriptions and theory* (pp. 147–170). John Benjamins.

Anthony, L. (2017). *AntConc (Version 3.5.0)* [Computer Software]. Retrieved from http://www.laurenceanthony.net/software

Ash, S. (2002). The distribution of a phonemic split in the Mid-Atlantic region: Yet more on short *a*. *University of Pennsylvania Working Papers in Linguistics*, *8*(3). Retrieved from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1741&context=pwpl

Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, *137*(5), 2823–2833.

Bayley, R. (2005). Second language acquisition and sociolinguistic variation. *Intercultural Communication Studies*, *14*(2), 1–15.

Beebe, L. M. (1980). Sociolinguistic variation and style shifting in second language acquisition. *Language Learning*, *30*(2), 433–447.

Beeching, K. (2015). Variability in native and non-native use of pragmatic markers: The wxample of *well* in role-play data. In K. Beeching & H. Woodfield (Eds.), *Researching Sociopragmatic Variability: Perspectives from Variational, Interlanguage and Contrastive Pragmatics* (pp. 174–197). Palgrave Macmillan.

Beeching, K. (2016). Like. In K. Beeching, *Pragmatic markers in British English: Meaning in social interaction* (pp. 126–155). Cambridge University Press.

Bucholz, M. (2011). White kids: Language, race, and styles of youth identity. Cambridge University Press.

Buchstaller, I. (2006). Social stereotypes, personality traits and regional perception displaced: Attitudes towards the "new" quotatives in the U.K. *Journal of Sociolinguistics, 10*(3), 362–381.

Buchstaller, I. (2014). *Quotatives: New trends and sociolinguistic implications*. Malden/Oxford: Wiley-Blackwell.

Butters, R. (1982). *Editor's note to Schourup*. *57*(2), 149.

Campbell-Kibler, K. (2007). Accent, (ing), and the social logic of listener perceptions. *American Speech*, *82*(1), 32–64.

Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change; Cambridge*, *21*(1), 135–156.

Campbell-Kibler, K. (2010). Sociolinguistics and perception. *Language and Linguistics Compass*, *4*(6), 377–389.

Campbell-Kibler, K. (2011). The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change*, *22*(3), 423–441.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47.

Chambers, J. K., Trudgill, P., & Schilling-Estes, N. (Eds.). (2002). *The Handbook of language variation and change*. Blackwell Wiley.

Chaudron, C., & Richards, J. C. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics, 7*(2), 113–127.

Cheshire, J., Kerswill, P., Fox, S., & Torgersen, E. (2011). Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics, 15*(2), 151–196.

Corder, S. P. (1981). Formal simplicity and functional simplification. In R. Anderson (Ed.), *New dimensions in second language acquisition research* (pp. 146–152). Rowley, MA: Newbury House.

Corrigan, K. P. (2015). "I always think of people here, you know, saying 'like' after every sentence." In C. Amador-Moreno, K. McCafferty, & E. Vaughan (Eds.), *The dynamics of discourse-pragmatic markers in Northern Irish English.* (pp. 37–64). John Benjamins.

Dailey-O'Cain, J. (2000). The sociolinguistic distribution of and attitudes toward focuser *like* and quotative *like*. *Journal of Sociolinguistics*, *4*(1), 60–80.

D'Arcy, A. (2005). *Like: Syntax and development* (Doctoral dissertation). Retrieved from https://search-proquest-com/docview/85634662

D'Arcy, A. (2007). Like and language ideology: Disentangling fact from fiction. *American Speech*, *82*(4), 386–419.

D'Arcy, A. (2011, 2014). Victoria English: Its development and current state [Research Grant]. *Social Sciences and Humanities Research Council of Canada (SSHRC). #410-2011-0219*.

D'Arcy, A. (2017). *Discourse-pragmatic variation in context: Eight hundred years of LIKE*. John Benjamins.

Darvin, R., & Norton, B. (2015). Identity and a model of investment in applied linguistics. *Annual Review of Applied Linguistics*, *35*, 36–56.

Davydova, J., & Buchstaller, I. (2015). Expanding the circle to learner English: Investigating quotative marking in a German student community. *American Speech*, *90*(4), 441.

De Costa, P. I. (2011). Using language ideology and positioning to broaden the SLA learner beliefs landscape: The case of an ESL learner from China. *System*, *39*(3), 347–358.

Dickerson, L. J. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly*, *9*(4), 401–407.

Dinkin, A. J. (2016). Variant-centered variation and the LIKE conspiracy. *Linguistic Variation; Amsterdam*, *16*(2), 221.

Diskin, C. (2013). Integration and identity: Acquisition of Irish-English by Polish and Chinese migrants in Dublin, Ireland. *Newcastle Working Papers in Linguistics*, *19*(1), 67–89.

Diskin, C. (2017). The use of the discourse-pragmatic marker 'like' by native and non-native speakers of English in Ireland. *Journal of Pragmatics*, *120*, 144–157.

Diskin, C., & Regan, I. (2015). Migratory experience and second language acquisition among Polish and Chinese migrants in Dublin, Ireland. In F. Forsberg Lundell, & I. Bartning (Eds.), *Cultural migrants and optimal language acquisition* (pp. 137–177). Bristol, Buffalo: Multilingual Matters.

Diskin, C., & Regan, V. (2017). The attitudes of recently-arrived Polish migrants to Irish English. *World Englishes; Oxford*, *36*(2), 191–207.

D'Onofrio, A. (2015). Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics*, *19*(2), 241–256.

Drager, K. (2016). Constructing style: phonetic variation in discursive functions of like. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: New methods and insights* (pp. 232–252). Cambridge University Press.

Drummond, R. (2011). Glottal variation in /t/ in non-native English speech: Patterns of acquisition. *English World-Wide, 32*(3), 280–308.

Durham, M., Haddican, B., Zweig, E., Johnson, D. E., Baker, Z., Cockeram, D., … Tyler, L. (2012). Constant Linguistic Effects in the Diffusion of Be Like. *Journal of English Linguistics, 40*(4), 316–337.

Eckert, P. (2005). *Variation, convention, and social meaning*. Paper presented at the the Annual Meeting of the Linguistic Society of America, Oakland, CA.

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics, 12*(4), 453–476.

Fleischman, S., & Yaguello, M. (2004). Discourse markers across languages: Evidence from English and French. In C.L. Moder, & A. Martinovic-Zic [Eds.], *Discourse across languages and cultures* (pp. 129-147). Amsterdam: John Benjamins.

Flowerdew, J., & Tauroza, S. (1995). The effect of discourse markers on second language Lecture Comprehension. *Studies in Second Language Acquisition, 17*(4), 435–458.

Fox Tree, J. E. (2007). Folk notions of *um* and *uh, you know*, and *like*. *Text & Talk, 27*(3), 297–314.

Fox Tree, J. E. (2010). Discourse Markers across Speakers and Settings. *Language and Linguistics Compass, 4*(5), 269–281.

Fuller, J. M. (2003a). Discourse marker use across speech contexts: A Comparison of native and non-native speaker performance. *Multilingua, 22*(2), 185–208.

Fuller, J. M. (2003b). The influence of speaker roles on discourse marker use. *Journal of Pragmatics, 35*(1), 23–45.

Fuller, J. M. (2003c). Use of the discourse marker *like* in interviews. *Journal of Sociolinguistics, 7*(3), 365–377.

Garrett, P., Williams, A., & Evans, B. (2005). Attitudinal data from New Zealand, Australia, the USA and UK about each other's Englishes: Recent changes or consequences of methodologies? *Multilingua, 24*(3), 211–235.

Gass, S. M., Spinner, P., & Behney, J. (2018). Salience in Second Language Acquisition and Related Fields. In S.M. Gass, P. Spinner, & J. Behney (Eds.), *Salience in Second Language Acquisition* (pp. 1–17). New York: Routledge.

Geeslin, K., & Long, A. (2014). *Sociolinguistics and Second Language Acquisition: Learning to Use Language in Context*. New York: Routledge.

Golato, A. (2000). An innovative German quotative for reporting on embodied actions: *Und ich so/und er so* and *I'm like/and he's like*. *Journal of Pragmatics, 32*(1), 29-54.

Hall-Lew, L., & Stephens, N. (2012). Country talk. *Journal of English Linguistics*, *40*(3), 256–

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4), 865–892.

Hays, P. R. (1992). Discourse markers and L2 acquisition. *Papers in Applied Linguistics-Michigan*, *7*, 24–34.

Hellermann, J., & Vergun, A. (2007). Language which is not taught: The discourse marker use of beginning adult learners of English. *Journal of Pragmatics*, *39*(1), 157–179.

Hesson, A., & Shellgren, M. (2015). Discourse marker LIKE in real time: Characterizing the time-course of sociolinguistic impression formation. *American Speech; Durham*, *90*(2), 154.

Jucker, A. H., & Smith, S. W. (1998). And people just you know like "wow": Discourse markers as negotiating strategies. In A.H. Jucker, & Y. Ziv (Eds.), *Discourse markers: Descriptions and theory*. John Benjamins.

Kern, J. (2014). Como in commute: The travels of a discourse marker across languages. *Studies in Hispanic and Lusophone Linguistics, 7*(2), 275-297.

Kerswill, P. (1994). *Dialects converging: Rural speech in urban Norway*. Oxford: Clarendon Press.

Kroskrity, P. (2004). Language ideologies. In *In A. Duranti (Ed.), A companion to linguistic anthropology* (pp. 496–517).

Labov, W. (1963). The social motivation of a sound change. *Word*, *19*(3), 273–309.

Labov, W. (1966). *The social stratification of English in New York City*. Washington: Center for Applied Linguistics.

Labov, W. (1972a). *Sociolinguistic patterns*. University of Pennsylvania Press.

Labov, W. (1972b). Some principles of linguistic methodology. *Language in Society*, *1*(1), 97–120.

Labov, W. (1984). Field methods of the Project on Linguistic Change and Variation. In J. Baugh & J. Scherzer (Eds.), *Language in Use: Readings in Sociolinguistics* (pp. 43–70). Englewood Cliffs: Prentice Hall.

Labov, W. (2001a). *Principles of linguistic change* (Vol. 2)*: Social factors*. Wiley-Blackwell.

Labov, W. (2001b). The anatomy of style-shifting. *In P. Eckert, & J. Rickford (Eds.), Style and sociolinguistic variation* (pp. 85–108). Cambridge University Press.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, *60*, 44–51.

Lavandera, B. R. (1978). Where does the sociolinguistic variable stop? *Language in Society, 7*(2), 171–182.

Lee, H. (1999). *Korean Americans as speakers of English: The acquisition of general and regional features* (Doctoral dissertation). Retrieved from http://search.proquest.com/llba/docview/85523699/D276BD378BE44470PQ/4

Levey, S. (2006). The sociolinguistic distribution of discourse marker "like" in preadolescent speech. *Multilingua*, *25*(4), 413–441.

Levin, H., & Gray, D. (1983). The Lecturer's OK. *American Speech*, *58*(3), 195–200.

Liamputtong, P. (2006). Motherhood and "moral career": Discourses of good motherhood among Southeast Asian immigrant women in Australia. *Qualitative Sociology*, *29*(1), 25–53.

Liao, S. (2009). Variation in the use of discourse markers by Chinese teaching assistants in the US. *Journal of Pragmatics*, *41*(7), 1313–1328.

Liu, B. (2016). Effect of L2 exposure: From a perspective of discourse markers. *Applied Linguistics Review*, *7*(1), 73–98.

Macaulay, R. (2002). Discourse Variation. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The Handbook of language variation and change* (pp. 283–305). Blackwell Wiley.

Maddeaux, R., & Dinkin, A. (2017). Is *like* like *like*?: Evaluating the same variant across multiple variables. *Linguistics Vanguard*, *3*(1).

Meehan, T. (1991). It's like, 'What's happening in the evolution of *like*?': A theory of grammaticalization. *Kansas Working Papers in Linguistics*, (16), 37–51.

Meyerhoff, M., & Schleef, E. (2012). Variation, contact and social indexicality in the acquisition of (ing) by teenage migrants. *Journal of Sociolinguistics*, *16*(3), 398–416.

Michigan State University (2018). *Planning and budgets: Enrollment* [Searchable data visualization]. Retrieved from: https://opb.msu.edu/functions/institution/datadigest/students/students-enrollment.html

Milroy, L. (2000). Social Network Analysis and Language Change: Introduction. *European Journal of English Studies*, *4*(3), 217–223.

Modaressi-Tehrani, Y. (1978). *A sociolinguistic analysis of modern Persian* (Doctoral dissertation). Retrieved from: https://search-proquest-com/llba /docview/302885658

Müller, S. (2005). *Discourse markers in native and non-native English discourse*. John Benjamins.

Nestor, N. (2013). The positional distribution of discourse *like*—a case study of young Poles in Ireland. In D. Singleton, V. Regan, & E. Debaene (Eds.), *Linguistic and Cultural Acquisition in a Migrant Community* (pp. 49–74). Bristol: Multilingual Matters.

Nestor, N., Chasaide, C. N., & Regan, V. (2012). Discourse 'like' and social identity – a case study of Poles in Ireland. In B. Migge, & M. Ní Chiosáin (Eds.), *New perspectives on Irish English.* (pp. 327–353). Amsterdam: John Benjamins.

Nestor, N., & Regan, V. (2015). The significance of age and place of residence in the positional distribution of discourse like in L2 speech. In C. P. Amador-Moreno, K. McCafferty, & E. Vaughan (Eds.), *Pragmatic markers in Irish English* (pp. 408–432). John Benjamins.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology, 18*(1), 62–85.

Niedzielski, N., & Preston, D. (2000). *Folk linguistics*. Mouton de Gruyter.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13.

Pichler, H. (2013). *The structure of discourse-pragmatic variation*. John Benjamins.

Pichler, H. (2016). Introduction: Discourse-pragmatic variation and change. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: New methods and insights* (pp. 1–18). Cambridge University Press.

Plonsky, L., Marsden, E., Crowther, D., Gass, S. M., & Spinner, P. (2019). A methodological synthesis and meta-analysis of judgment tasks in second language research. *Second Language Research*, 1–39.

Preston, D. R. (2013). Language with an attitude. In J.K. Chambers, & Natalie Schilling (Eds.), *The Handbook of language variation and change* (pp. 157–182). Wiley Blackwell.

Preston, D. R., & Bayley, R. (Eds.). (1996). *Second language acquisition and linguistic variation*. John Benjamins.

Regan, V. (2010). Sociolinguistic competence, variation patterns and identity construction in L2 and multilingual speakers. *EUROSLA Yearbook, 10*(1), 21–37.

Regan, V., Howard, M., & Lemée, I. (2009). *The Acquisition of Sociolinguistic Competence in a Study Abroad Context*. Multilingual Matters.

Reyes, A. (2007). *Language, identity, and stereotype among Southeast Asian American youth : The Other Asian*. New Jersey: Lawrence Erlbaum.

Romaine, S., & Lange, D. (1991). The use of "like" as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech*, *66*(3), 227–279.

Romero Trillo, J. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, *34*(6), 769–784.

Russell, B., Perkins, J., & Grinnell, H. (2008). Interviewees' overuse of the word "like" and hesitations: Effects in simulated hiring decisions. *Psychological Reports*, *102*(1), 111–118.

Sankoff, G., Thibault, P., Nagy, N., Blondeau, H., Fonollosa, M.-O., & Gagnon, L. (1997). Variation in the use of discourse markers in a language contact situation. *Language Variation and Change*, *9*(2), 191–217.

Schleef, E. (2008). The "Lecturer's Ok" revisited: Changing discourse conventions and the influence of academic division. *American Speech*, *83*(1), 62–84.

Schleef, E. (2013). Developmental sociolinguistics and the acquisition of T-glottaling by immigrant teenagers in London. In G. De Vogelaer, & M. Katerbow (Eds.), *Acquiring sociolinguistic variation* (pp. 305–341). John Benjamins P.

Schourup, L. (1985). *Common discourse particles in English conversation*. New York: Routledge.

Schweinberger, M. (2011). *The discourse marker LIKE: a corpus-based analysis of selected varieties of English* (Doctoral dissertation). Retrieved from: https://www.researchgate.net/publication/329202381_The_discourse_marker_LIKE_a_corpus-based_analysis_of_selected_varieties_of_English

Schweinberger, M. (2012). The discourse marker LIKE in Irish English. In B. Migge, & M. Ní Chiosáin (Eds.), *New perspectives on Irish English.* (pp. 179–201). Amsterdam: John Benjamins.

Sharma, D. (forthcoming). *Biographical indexicality: Personal history as a frame of reference for social meaning in variation*.

Sharma, D. (2016). Changing ethnicities: The evolving speech styles of Punjabi Londoners. *In H. S. Alim, J. R. Rickford, & A. F. Ball (Eds.), Raciolinguistics: how language shapes our ideas about race* (pp. 221–240). Oxford: Oxford University Press.

Siemund, P., Maier, G., & Schweinberger, M. (2009). Towards a more fine-grained analysis of the areal distributions of non-standard features of English. In H. Paulasto and E. Penttilä

(Eds.), *Language contacts meet English dialects: Studies in honor of Markku Filppula* (pp. 19–46). Cambridge: Cambridge Scholars Publishing.

Silverstein, M. (1979). Language structure and linguistic ideology. In R. Clyne, W. Hanks, & C. Hofbauer (Eds.), *The Elements: A Parasession on linguistic units and levels* (pp. 193–247).

Singler, J. V. (2001). Why you can't do a VARBRUL study of quotatives and what such a study can show us. *University of Pennsylvania Working Papers in Linguistics*, *7*(3), 257–278.

Spinner, P., Behney, J., & Gass, S. M. (2018). Afterword: The role of salience in second language research. In S.M. Gass, P. Spinner, & J. Behney (Eds.), *Salience in Second Language Acquisition* (pp. 291–297). New York: Routledge.

Spinner, P., & Gass, S. (2019). *Using judgments in second language acquisition research*. New York: Routledge.

Svartvik, J. (1980). Well in conversation. In S. Greenbaum, G. Leech, & J. Svartvik (Eds.), *Studies in English Linguistics for Randolph Quirk* (pp. 167–177). London: Longman.

Tagliamonte, S. (2016). *Teen talk: The language of adolescents*. Cambridge University Press.

Tagliamonte, S. A., & D'Arcy, A. (2007). Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change; Cambridge*, *19*(2), 199–217.

Tagliamonte, S., & D'Arcy, A. (2004). He's like, she's like: The quotative system in Canadian youth. *Journal of Sociolinguistics*, *8*(4), 493–514.

Tagliamonte, S., & Hudson, R. (1999). *Be like* et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics*, *3*(2), 147–172.

Taguchi, N., & Roever, C. (2017). *Second language pragmatics*. Oxford: Oxford University Press.

Talmy, S. (2010). Qualitative interviews in applied linguistics: From research instrument to social practice. *Annual Review of Applied Linguistics*, *30*, 128–148.

Talmy, S., & Richards, K. (2011). Theorizing qualitative research interviews in applied linguistics. *Applied Linguistics (Oxford)*, *32*(1), 1–5.

The Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a Multilingual World. *The Modern Language Journal*, *100*(S1), 19–47.

Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, *1*(2), 179–195.

Truesdale, S., & Meyerhoff, M. (2015). Acquiring some *like*-ness to others: how some Polish teenagers acquire the Scottish pragmatics of *like*. *Te Reo*, *58*, 3.

Tyler, A. (1992). Discourse structure and the perception of incoherence in international teaching assistants' spoken discourse. *TESOL Quarterly*, *26*(4), 713–729.

Tyler, A. E., Jefferies, A. A., & Davies, C. E. (1988). The effect of discourse structuring devices on listener perceptions of coherence in non-native university teacher's spoken discourse. *World Englishes*, *7*(2), 101–110.

Underhill, R. (1988). *Like* is, *like*, focus. *American Speech*, *63*(3), 234–246.

Wagner, S., Mason, A., Nesbitt, M., Pevan, E., & Savage, M. (2016). Reversal and re-organization of the Northern Cities Shift in Michigan. *University of Pennsylvania Working Papers in Linguistics*, *22*(2), 171–179.

Williams, J. (1992). Planning, discourse marking, and the comprehensibility of international teaching assistants. *TESOL Quarterly*, *26*(4), 693–711.

Winter, J. (2002). Discourse quotatives in Australian English: Adolescents performing voices. *Australian Journal of Linguistics*, *22*(1), 5–21.

Woolard, K. A., & Schieffelin, B. B. (1994). Language ideology. *Annual Review of Anthropology*, *23*(1), 55–82.

Young, R. (1988). *Approaches to variation in interlanguage: Plural marking in the speech of Chinese learners of English* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia.