The Development of Disability and Foreignness Concepts: A Comparative Approach

by

Netanel Y. Weinstein

A dissertation accepted and approved in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in Psychology

Dissertation Committee:
Dare A. Baldwin, Chair
Sara D. Hodges, Core Member
Melissa Baese-Berk, Core Member
Melissa A. Redford, Institutional Representative

University of Oregon

Summer 2024

DISSERTATION ABSTRACT

Netanel Y. Weinstein

Doctor of Philosophy in Psychology

Title: The Development of Disability and Foreignness Concepts: A Comparative Approach

Human cognition often displays a tendency to "see beyond" the available perceptual input. Although inferences indicative of such seeing-beyond tendencies are fundamental to efficiency in human cognition, they may also be associated with the expression of prejudice towards stigmatized others. In this dissertation, we systematically compared college students' and children's inductive generalization tendencies regarding two stigmatized social categories: foreignness and disability. Since such cues may be apparent both in speech (i.e., a foreign accent; a speech disorder) and appearance (i.e., foreign garb; a wheelchair), directly juxtaposing children and college students' reactions to such cues may be particularly informative regarding the development of foreignness and disability concepts alike.

In a first study, we compared 180 North American college-aged students' and 163 young children's (M_{age} = 5.75) explicit assessments of a) three speech categories (neurotypical American English - L1, Spanish English - L2, and American English with Autism Spectrum Disorder - ASD), and b) four illustration categories (children whose appearance was: able-bodied typical North American appearance; able-bodied foreign appearance; typical North American wheelchair-bound with signs of contracture in the wrist and torticollis in the neck; and typical North American amputee appearance) along several key dimensions (i.e., foreignness, dependence, competence, interest in friendship and comprehensibility for speech). To further explore developmental change in inductive generalization tendencies, in study 2, we assessed 130 college-aged students' and 143 North American children's (M_{age} = 5.3) associations

between speech variability and visual appearance. Specifically, participants listened to one of three speech conditions (L1, L2, ASD) while looking at two illustrations side-by-side (one of a typical American child, the other depicting a foreign child or a child with a disability) and were asked to select the child who was talking.

Across both studies, college students, but not children, appeared to associate the variability they detected in the ASD speech with a latent disability concept in a similar manner to which both samples associated L2 speech with foreignness. Nevertheless, there was an emerging age-related increase in this tendency for children as well, particularly for those with advanced metacognition. Furthermore, whereas college students were biased against ASD speakers but showed a prosocial bias towards images depicting physical disabilities (particularly amputees), children were biased against wheelchair (and foreign) images but showed no bias in the case of ASD speech. This work advances our understanding of complex ways in which conceptual representations of the social world relate to the expression of prejudice, and how such relationships may change developmentally. Our findings also hold potential to inform development of empirically-oriented interventions to reduce the expression of prejudice in childhood and across the lifespan.

CURRICULUM VITAE

NAME OF AUTHOR: Netanel Y. Weinstein

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
The New School for Social Research, New York NY
Yeshiva University, New York NY

DEGREES AWARDED:

Doctor of Philosophy, 2024, University of Oregon
Master of Arts, Psychology, 2017, The New School for Social Research
Bachelor of Arts, Psychology, 2014, Yeshiva University

AREAS OF SPECIAL INTEREST:

Developmental Psychology
Cultural Psychology
Sociolinguistics

PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, University of Oregon Department of Psychology, 2017-2024

GRANTS, AWARDS, AND HONORS:

(2024) University of Oregon Graduate Education Committee Research Award ($500)
(2022) University of Oregon Graduate Education Committee Research Award ($500)
(2019) Gorilla Experiment Builder Grant ($500)

PUBLICATIONS:

Baldwin, D., Measelle, J., Gallivan, L., Sanchirico, A., **Weinstein, N**., Bala, A., Chan, K, Gallant, J. H., Borath, M., Kroeun, H., Prak, S., Wieringa, F., Green, T. J., Whitfield, K. C. (in press @ Developmental Psychology). Language processing in breastfed infants at risk of thiamine deficiency benefits from maternal thiamine supplementation.

Weinstein, N., Helienbach, S., Oshiro, M., and Baldwin, D. (forthcoming @ *Cognitive Development*). The Development of Disability and Foreignness Concepts: A Comparative Approach.

Everett, Y., Lightcap, A., O'Brien, J., **Weinstein, N**., and Zalewski, M. (2024). Integrating Dialectical Behavior Therapy Skills and Parent Training for Dually Dysregulated Parents and Children: An Idiographic Case Study. *Cognitive and Behavioral Practice*. https://doi.org/10.1016/j.cbpra.2024.04.002

Weinstein, N. Y., and Baldwin, D. A. (2023). The many faces of moralized self-control: Puritanical morality is not reducible to cooperation concerns. *Behavioral and Brain Sciences*, *46*, e320. https://doi.org/10.1017/S0140525X23000419

Weinstein, N., and Baldwin, D. (2022). Reification of infant-directed speech? Exploring assumptions shaping infant-directed speech research. *Culture and Psychology*. https://doi.org/10.1177/1354067X22114768

Weinstein, N., Whitmore, L. B., and Mills, K. L. (2022). Individual Differences in Mentalizing Tendencies. *Collabra: Psychology*, *8*(1), 37602. https://doi.org/10.1525/collabra.37602

Baldwin, D. A., Measelle, J., Gallivan, L., Sanchirico, A., **Weinstein, N**., Kroeun, H., ... and Whitfield, K. (2021). Maternal Thiamine Supplementation Promotes Infants' Language Processing at 24 Weeks. *Current Developments in Nutrition*, *5*(Supplement_2), 892-892. https://doi.org/10.1093/cdn/nzab049_005

Weinstein, N., and Baldwin, D. (2020). Causal learning in CTC: Adaptive and collaborative. *Behavioral and Brain Sciences*, *43*. DOI:10.1017/S0140525X20000199

ACKNOWLEDGMENTS

I would first like to thank professors Melissa Redford and Àngels Llanes for generously sharing the speech samples and Jesse Springer for producing the illustrations used in this dissertation. I would also like to thank the Living Lab at the Oregon Museum of Science and Industry (OMSI) for supporting the data collection with children. I am particularly grateful to Annie Douglass and Chris Cardiel for providing valuable feedback regarding pragmatic and ethical aspects of this research.

I am also grateful for my colleagues and friends in the Acquiring Minds Lab, past and present. I am particularly grateful to Dianna DeWald-Daschel for establishing a research collaboration with the Living Lab at OMSI and for supporting my efforts in collaborating with them. I would also like to thank my undergraduate research assistants and collaborators, Samuel Heilenbach and Megan Oshiro, for their data collection efforts and support in undertaking the analysis. I am especially grateful for my advisor, Professor Dare Baldwin for seven years of mentorship. Dr. Baldwin's endless curiosity, intellectual depth and meticulous attention to detail have been a great inspiration. For me, Dare exemplifies the very best of academic scholarship.

Finally, I would like to thank my wife, Yael, for hours of conversation and debate about the merits of my work and my children Leora, Leba and Zev for all the love and support I could wish for.

TABLE OF CONTENTS

Chapter Page

Chapter Page

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

Table Page

Table Page

Table Page

Table Page

Table Page

Table Page

Table Page

CHAPTER 1

INTRODUCTION

1.1. Overview

The social world is infinitely complex. At any given moment, a person can be categorized by a wide range of features including (but not limited to) their age, gender, ethnicity, kinship, religion, occupation, accent, competencies, interests, and ability status. Social categorization thus exemplifies a basic inductive challenge that characterizes categorization efforts more generally (Markman, 1989; Quine, 1960), given the potentially infinite range of categorization schemes that are theoretically possible for any collection of items.

To navigate this complexity, even young children appear to construct intuitive theories regarding the structure and nature of the social world (Banaji and Gelman, 2013). For example, children may intuit that the social world can be divided into discrete groups that share certain characteristic features. Thus, they might expect people who dress a certain way or live in certain types of homes to also talk a certain way (Hirschfeld and Gelman, 1997). Beyond mere associationism, intuitive/folk sociological theories often attribute observed group commonalities to (presumed) essential features of the group rather than to other contextual attributions (Hirschfeld, 1996). Indeed, the ability to detect salient social variation (i.e., in speech, behavior or appearance) and attribute such variability to a latent group concept (or proto concept) is argued to reflect a core feature of human cognition that is functional already in infancy (e.g., Hirschfeld, 2003; Spelke and Kinzler, 2007).

Whereas the bulk of research on children's understanding of the social world has focused on perceptions of social groups such as race (Mandalaywala et al., 2019), gender (Gelman and Taylor, 2014) and nationality (Hussak and Cimpian, 2019), considerably less theoretical and

empirical work has been devoted to exploring children's understanding of another sort of variability that is often also perceptually apparent - disability. This gap is somewhat surprising since perceptions of competence (which are likely central to perceptions of disability) are considered to be fundamental to the person-perception process and, in fact, have been described as "universal" (e.g., Cuddy, Fiske and Glick, 2008; Halkias and Diamantopoulos, 2020). Based on these accounts, there is reason to expect that children would be perceptually attuned to meaningful variability in their peers' competencies in a similar manner to which they might be attuned to their gender, race, or ethnolinguistic group membership.

Furthermore, beyond merely noting variability in competencies, we might also expect that children would be inclined to categorize and explain the origins of this variability. Since continuous perceptual variability is often perceived in categorical terms (e.g., perceiving discrete colors in a continuous electromagnetic spectrum; Goldstone and Hendrickson, 2010), we might expect children to attribute perceptions of reduced competence to a latent categorical disability concept in a similar manner to which they might attribute variability in skin tone to a latent race concept, particularly if they are familiar with linguistic conventional labels denoting such concepts. In turn, these categorical representations may be associated with a range of inductive generalizations about the individual that extend beyond the perceived variability (Leslie, 2017).

In the context of research on perceptions of individuals with a disability, such inductive generalization tendencies are sometimes described as "spreading" (Wright, 1964;1983). To illustrate, Liesener and Mills (1999) found that nondisabled college students tended to provide instructions to physically disabled adult targets in a similar manner to the way they assisted children, using higher pitch and more words. Critically, the participants exhibited this tendency even when the targets with a disability were framed as managers (and therefore, by implication,

not dependent or cognitively impaired). In this study, non-disabled college students appeared to have drawn a spreading inference that the physical disabilities they observed in adults were indicative of cognitive disability as well, engendering a need for modification to simplify communicative messages. Thus, physical disability appeared to act as a conceptual anchor for the constructive person-perception process (Asch, 1946), resulting in participants' relating to individuals with a disability as *fundamentally* disabled (Dunn, 2015).

The tendency for a particular kind of perceptual cue, such as a cue to physical disability, to dominate the person-perception process potentially bears some relation to Goffman's (1963) influential discussion of stigma. In that framework, when perceptually salient cues also carry a social stigma within a given social context (e.g., certain skin tones, accents, disabilities), observers may focus on the stigmatized social category to which the individual belongs, thus devaluing the personal history of the individuals' actual identity. In such cases, "the stigmatized individual may become "invisible" because of the visibility and social meaning of the stigmata" (Loury, 2005:2; see also Loury, 2021).

Yet in contrast to the intrinsically prejudicial connotation of stigma, spreading is not necessarily grounded in the existence of prejudice towards individuals with a disability. Thus, raising one's voice when speaking to a visually impaired person or treating someone who is physically disabled as though they are intellectually disabled need not necessarily stem from prejudicial attitudes toward individuals with a disability (Liesener and Mills 1999). In fact, behavioral modifications associated with disability-related spreading may often arise from a sympathetic motivation to provide greater assistance to others perceived as in need, rather than from a negative bias or prejudice. Indeed, disability-related spreading has been explicitly

contrasted with the concept of stigma to emphasize how generalizing processes anchored on a disability cue are distinct from the expression of negative attitudes (Liesener and Mills, 1999).

Despite the theoretical framing of spreading as orthogonal to stigma (in at least some cases), there is limited empirical research directly exploring this issue, particularly from a developmental perspective. In fact, only a handful of studies (e.g., Smith and Williams, 2001; Wheeler and Tharpe, 2020) have even explored spreading tendencies in childhood and these studies do not typically assess relations between these tendencies and expression of prejudice (though see Werner, Peretz, and Roth, 2015). Furthermore, it is likely that, at least under certain circumstances, spreading may also be directly tied to prejudice. For example, Parker, Monteith and South (2020) found that the tendency to associate individuals with developmental disabilities (ASD and Down's syndrome) with enhanced dependence was positively associated with the tendency to dehumanize these individuals (i.e., to negate the premise that features considered to be unique to humans such as refinement, rationality, and self-restraint were characteristic of individuals with ASD and Down's; Haslam, 2006). Thus, the limited empirical evidence regarding a) the relationship between spreading tendencies and the expression of prejudice, and b) the possibility that spreading may sometimes also be directly tied to prejudice, calls for renewed investigation of these issues, particularly from a developmental perspective.

In sum, although normative inductive processes that enable perceivers to utilize salient perceptual cues to categorize people quickly and efficiently are likely crucial for survival (i.e., by recognizing threat; Krienert, Walsh, and Acquaviva, 2018), the tendency to draw generalized social inferences about others based on minimal exposure may also contribute to undesirable social outcomes such as social rejection and bias (Sasson et al., 2017). Due to the centrality of inductive generalization to social cognition (and cognition more generally; Markman, 1989;

Baldwin, Markman and Melartin, 1993; Griffiths et al., 2010) and the possibility that inductive processes may be related to stigmatizing processes (Leslie, 2017), there is an urgent need to advance our understanding of the development of inductive generalizations in the case of stigmatized social categories.

1.2.Inductive Generalizations - A Comparative Approach

In this dissertation, I take a comparative approach to these issues by systematically examining children's inductive generalization tendencies regarding two stigmatized social categories: foreignness and disability. Since such cues may be apparent both in speech (i.e., a foreign accent; a speech disorder) and appearance (i.e., foreign garb; a wheelchair), directly juxtaposing children's reactions to such cues may be particularly informative regarding children's attunement to the divergent social significance of these different sorts of stigmatized cues. In turn, enhancing our understanding of this level of attunement may support efforts to foster more positive social relationships and reduce early emerging signs of prejudice.

The comparative approach taken here is consistent with recent theoretical accounts of children's perceptions of disability suggesting that non-disabled children may view disability as a "special kind of out-group" (Babik and Gardner, 2021:1). According to this recent proposal, the expression of prejudice against individuals with a disability may be rooted in basic social psychological processes motivating people to protect and promote the self by derogating members of an out-group (e.g., Tajfel and Turner, 1979). Beyond accounting for the expression of prejudice, framing disability as an out-group also implies that common inductive processes may undergird perceptions of ethnolinguistic variability and disability alike. Specifically, just as children may anchor upon perceptually available cues such as non-native accents to draw

broader inferences regarding the sort of (outgroup) person the speaker is, so too might they view a perceived impairment as indicative of broader, unseen aspects of the person.

Along these lines, Wheeler and Tharpe (2020) found that children as young as 6 years of age perceived peers who were wearing hearing aids as less physically competent (e.g., less good at swimming/climbing/running) than those without hearing aids (the so-called "hearing-aid effect"). Similarly, Smith and Williams (2001) found that preschoolers expected children who were wheel-chair-bound to have reduced cognitive competence (e.g., reduced skill at puzzles/reading). In these studies, the detection of a disability in one domain (hearing, walking) led to the expectation that the affected individual would exhibit disability in another, unrelated domain (physical, cognitive).

In what follows, I provide an overview of previous research assessing children's attunement to perceptual variability attributable to foreignness and disability. In this overview, I focus primarily on children's perceptions of speech variability. This approach is justified considering the rich history of scholars from a range of fields who have described speech as particularly central to perceptions of personality (Sapir, 1927), group affiliation (Kinzler and Liberman, 2017), and disability status (Grandin and Scariano, 1986) alike.

Along these lines, writing in 1927, Edward Sapir surmised that "the voice is in some way a symbolic index of the total personality" (p 896). Relatedly, reflecting on her own social experiences as an autistic child, Temple Grandin noted that "My voice was flat with little inflection and no rhythm. That alone stamped me as different" (Grandin and Scariano, 1986, p. 21). More recently, Kinzler and Liberman (2017:3753) suggested that "infants make a constellation of inferences based on a speaker's language or accent, which could hang together as a larger conceptual representation of group membership defined by language".

Throughout our discussion of the extant literature, we point to a series of open questions that we hope to address in the present dissertation. We then describe the dissertation research and explain how the methods adopted can address the open questions stemming from this literature.

1.3.Children's Perceptual Attunement to Foreignness Cues in Speech

As previously noted, several researchers investigating young children's social categorization processes have argued that speech operates as a particularly salient source of perceptual social information (Kinzler, 2021). For example, Liberman, Woodward and Kinzler (2017) have argued that infants view language as a meaningful social marker and use language to draw inferences about third-party social relationships and food preferences (Liberman et al., 2016). The common thread throughout this work is that speech variability is said to act as a particularly potent source of information that children rely upon to categorize the social environment into distinct groups.

Indeed, several studies have found that children as young as 4 years of age reliably relate language differences to social differences. For example, North American 4-year-olds consistently mapped foreign/non-Western items (such as an igloo) to a foreign language with which they were unfamiliar (Portuguese) (Hirschfeld and Gelman, 1997). In a similar vein, Wagner, Clopper and Pate (2014) extended these findings in a sample of 5–6-year-olds with a Second-Language dialect (Maharashtran variety of Indian English). The emerging conclusion from these studies is that preschoolers reliably relate speech differences to group differences and expect different groups to share distinct ways of speaking (see also Weatherhead, White and Friedman, 2016).

Furthermore, the claim that speech is particularly important for social cognition is based on findings with older participants as well. For example, Rakić, Steffens and Mummendey (2011) found that while German college students tended to spontaneously categorize speakers based on both accents and appearance, accents appeared to act as the more meaningful cue to ethnic category membership. In fact, the notion that detection of speech variability plays a vital role in social categorization is a fundamental premise of several research programs in fields such as sociolinguistics and linguistic anthropology as well (Campbell-Kibler, 2007; Eckert, 2008; Labov, 1964).

On the other hand, there is some ambiguity regarding children's (proposed) precocious sociolinguistic awareness. For example, some findings indicate that accent categories may not represent a salient social distinction to preschool-aged children in the same way as gender or age (e.g., Creel, 2012; Creel and Jiménez, 2012). Specifically, Creel (2018) found that North American children (ages 3-7) showed limited ability to associate a Second-Language dialect of English (Dutch-English) with "not being from here". Furthermore, although their overall sample did perform above chance, this effect was largely driven by the older children. In fact, Barrett (2007) has argued that children do not begin to map language differences onto national or cultural differences until the age of 6.

Due to the tension between early-competence accounts depicting relatively stable precocious sociolinguistic awareness from infancy and protracted accounts of developing sociolinguistic awareness, more research is needed to adjudicate between these divergent accounts.

1.4.Children's Perceptual Attunement to Disability Cues in Speech

In parallel to the ongoing research on children's attunement to (and reasoning about) speech variability in the case of accents and dialects, several studies have also explored children's perceptions of speech variability that is attributable to a disorder. Thus, previous studies have found that speakers with cleft lip/palate (Lee, Gibbon, and Spivey, 2017). articulation disorders (Hall, 1991; Freeby and Madison, 1989), voice disorders (Lass et al., 1991), nasal resonance disorders (Blood and Hyman, 1977), motor speech disorders (Lass et al., 1993), and those who stutter (Franck et al., 2003) were all generally perceived in a negative manner by other children.

Indeed, children with various forms of Speech Disorder often experience social challenges associated with their speech (Krueger, 2019). These challenges may contribute to fewer opportunities to build strong relationships or friendships with peers, an important factor in the development of self-confidence, socio-emotional and socio-cognitive skills, which may in turn have a negative impact on psychosocial functioning, education, and employment (Bettens et al., 2020). Shedding light on the foundations of negative reactions to non-standard speech is therefore critically important for the developmental well-being of children with speech disorder – an estimated 8% of all children in the US (Lindsay and Strand, 2016).

However, previous research on these issues was primarily based on responses provided by children aged 7 and older, in contrast to the literature on accents and dialects, which has investigated younger children as well. Important gaps regarding the onset of prejudicial reactions to speech disorders, therefore, remain. Furthermore, since previous research has typically presented children with highly salient disordered speech, the extent to which children are perceptually attuned to more subtle indicators of a speech disorder remains unclear. This gap is important to address since even subtle indicators of disordered speech may be highly

consequential for the social reception of speakers, to the extent that peers indeed detect these subtleties.

To illustrate, Redford, Kapatsinski and Cornell-Fabiano (2018) presented college student participants with speech excerpts produced by a) neurotypical children on the one hand, and b) children with autism spectrum disorder exhibiting unusual speech patterns on the other hand, and asked participants to determine whether the speech they were hearing was "typical" or "disordered". Participants also provided likability ratings for the two groups of children based solely on these speech excerpts. Redford and colleagues found that two-second excerpts were sufficient to elicit significantly greater identification of disordered speech with the ASD sample, as well as significantly lower likability ratings for the ASD speakers. Whether children would exhibit similar levels of attunement (as well as similar aversive reactions) as did college student participants in that study is currently unknown.

1.5.Limitations of Previous Research

Even though previous theoretical accounts (e.g., Babik and Gardner, 2021) of children's perceptions of disabilities have pointed to analogous processes operating to guide perceptions of foreignness, no empirical work. as yet. has directly investigated this proposal. Concretely, it is unclear the extent to which children exhibit similar attunement to salient cues typically associated with a disability (e.g., a speech disorder or wheelchair) versus cues that are typically associated with other kinds of group membership (e.g., accent or foreign garb). Furthermore, beyond perceptual attunement, it is unclear the extent to which young children reliably attribute certain percepts (e.g., speech disorder, wheelchair) to a latent and inductively rich disability concept in a similar manner to which they might attribute other percepts (e.g., accent, foreign garb) to a foreignness concept.

In fact, previous research has not determined the extent to which children distinguish between a speech disorder and a foreign accent when drawing generalizing inferences about speakers with a speech disorder or accent. Furthermore, in a rare study involving specialists in voice therapy (ENT, phoneticians, speech therapists), Ghio et al., (2013) found that, at least under certain circumstances, even expert listeners may have difficulty distinguishing between an accent and speech disorder. This suggests that young children are likely to have difficulty in differentiating between foreign accent and speech disorder, which would plausibly affect the possibility of generalizing inferences.

The comparative approach taken here is promising as it sheds light on children's emerging attunement to (and reasoning about) salient social cues related to disability and foreignness alike. This approach makes it possible to contrast children's responses to these various saliencies, thus contributing greater insight into the developmental emergence of these divergent social concepts.

Assessing developmental changes in the degree of children's acuity in detecting perceptually salient cues and sorting them into coherent explanatory frameworks is vital to understanding the emergence of prejudicial attitudes towards individuals with a disability. For example, previous research indicated that adults often harbor greater negative attitudes towards individuals with cerebral palsy in comparison to individuals with an amputation (Gordon, Minnes and Holden, 1990). The extent to which children follow a similar pattern depends on their ability to detect certain postural characteristics that are often associated with cerebral palsy but not an amputation (e.g., contracture in the wrist and torticollis in the neck). Previous research has documented that young children (ages 3-6) show greater attunement to (and awareness of) the consequences of physical disabilities in comparison to other disabilities such

as visual/hearing impairments or Down's syndrome (Diamond and Hestenes, 1996). However, no previous studies have directly compared children's responses to an amputation versus a physical disability characterized by a constellation of perceived impairments such as cerebral palsy.

Beyond shedding light on the emergence of disability-related prejudice, assessing participants' attunement to (and understanding of) disability cues is informative regarding the construct of disability spread. Namely, previous accounts of disability spread (Liesener and Mills, 1999; Wright, 1983) have typically highlighted inaccurate inductive generalization processes (e.g., speaking to a visually impaired person as though they were deaf; speaking to a manager in a wheelchair as though they were cognitively disabled). Due to this previous emphasis on inaccuracy, it is unclear whether reasonably accurate inductive generalizations related to perceptions of disability constitute disability spread as well. Furthermore, the extent to which individuals consider the sort of impairment they encounter (e.g., amputation vs. severe cerebral palsy) when drawing generalizing inferences regarding the target (e.g., by engaging in disability spread regarding a target with severe cerebral palsy but not an amputee), is also not clear.

In sum, this dissertation aims to address a range of remaining questions related to children's developing attunement to cues to both disability and foreignness in speech and appearance, the inductive generalizations associated with this attunement, and possible relations between these generalizations and expressions of prejudice.

1.6. Overview of the Proposed Dissertation

1.6.1. Overarching Goals

The overarching goals of this dissertation were to shed light on 1) children's developing tendency to discriminate between foreignness and disability cues, 2) differential generalizations stemming from these discriminations, and 3) expressions of prejudice towards foreign individuals and individuals with a disability. We therefore explored these factors in the studies described below.

1.6.2. An Explicit Lexical Approach

In a first study, we compared North American children's (n = 163, M_{age} = 5.75, SD = 1.6, range = 3:9) and college-aged students' (n = 180, M_{age} = 18.9 , SD = 4.4, range = 18:28) explicit assessments of three speech categories (children whose speech was either neurotypical North American-accented English – L1; neurotypical Spanish-accented English – L2; and North American-accented English produced by children with autism spectrum disorder - ASD) along several key dimensions (i.e., foreignness, dependence, competence, interest in friendship and comprehensibility for speech). As well, we compared the same individuals' assessment of four illustration categories (children whose appearance was: able-bodied typical North American appearance; able-bodied foreign appearance; typical North American wheelchair-bound with signs of contracture in the wrist and torticollis in the neck; and typical North American amputee appearance) along the same set of dimensions.

This approach facilitated exploration of age-related changes in the selectivity of disability and foreignness concepts from early childhood to emerging adulthood.

To illustrate, asking participants whether L2 and ASD speakers were "from another country far away" was informative regarding children's developing conceptual differentiation between two very different sorts of perceptual variability (i.e., accent and speech disorder). Furthermore, comparing the extent to which participants assumed that L2 and ASD speakers

"need a lot of help to do things" or are "good at running" was informative of the extent to which children draw different sorts of inductive generalizations about speakers with a foreign accent versus a speech disorder. Namely, associating high dependence and reduced physical competence with ASD speakers, but not L2 speakers, would indicate the operation of disability spread (as it is typically conceptualized). If, instead, participants associated both ASD and L2 speech with high dependence and reduced physical competence, this would indicate a more generalized tendency to project dysfunction on individuals exhibiting any form of atypical speech (and thus, not disability spread, per se). Finally, asking participants whether they were interested in being friends with the speakers allowed us to assess prejudicial attitudes (as well as the extent to which spreading tendencies were related to the expression of prejudice).

1.6.3. An Implicit Associative Approach

The method used in study 1 relied on explicit judgments of countries (i.e., "America" versus "another country far away") which young children may struggle with (Wagner, Greene-Havas, and Gillespie, 2010). Thus, a method that does not rely on such knowledge, such as measuring associations between speech and appearance (e.g., Hirschfeld and Gelman, 1997; Wagner, Clopper and Pate, 2014) was warranted to further explore young children's intuitions regarding the social significance of variation in speech and appearance. To this end, in study 2, we assessed North American children's ($n = 143$, $M_{age} = 5.3$, $SD = 1.8$, range = 3:11) and college-aged students' ($n = 130$, $M_{age} = 19.9$, $SD = 1$, range = 18:26) associations between speech variability and visual appearance. Specifically, participants listened to one of three speech conditions (L1, L2, ASD) while looking at two illustrations side-by-side (one of a typical American child, the other depicting a foreign child or a child with a disability) and were asked to select the child who was talking.

This approach facilitated more direct exploration of participants' inductive generalizations without relying upon explicit indexical judgments. Thus, the extent to which participants would selectively associate ASD speech with a physical disability but not foreign garb (and vice versa: L2 speech with foreign garb but not physical disability) was informative regarding the emergence of disability spread and its relation to inductive generalization in the case of foreignness concepts. Furthermore, comparing the extent to which participants would selectively associate ASD speech with images depicting a severe case of cerebral palsy to a greater extent than images depicting an amputee was also informative regarding the way factors such as severity and kind of disability relate to disability spreading tendencies.

1.6.4. Metacognition Assessment

To map speech variability onto a social category, children must have some ability to reflect upon the speech they are hearing and contemplate why speakers sound the way that they do. For example, they need to recognize that some speech is more difficult for them to understand than others and explain why that may be the case. While such reflective processes may seem relatively straightforward, a long line of ongoing research suggests that younger children experience significant difficulty accurately reflecting upon their own cognitive processes, such as whether they comprehend something (e.g., Markman 1977;1979; Gascoine, Higgins and Wall, 2017).

Given the key conceptual relationship between metacognition and sociolinguistic awareness, it is somewhat surprising that previous research has ignored this relationship and instead focused on other domain-general cognitive abilities, such as working memory, that also are likely related to variability in children's performance on tasks measuring sociolinguistic awareness. For example, children participating in sociolinguistic categorization tasks must be

able to simultaneously remember the features of a speech stream and link that speech to a specific category. Since such tasks are considerably taxing from a working memory perspective, younger children's apparent ignorance of the social significance of sociolinguistic variability (e.g., Labov, 1964), may alternatively be seen as an artefact of the experimental design (see Jones et al., 2016 for discussion).

In this dissertation, we introduce a novel methodological innovation that may shed further light on the relationship between domain-general cognitive capacities and the nature of age-related change in performance on sociolinguistic tasks. Specifically, in all studies, we measured participants' ability to accurately report whether it was "easy" for them to understand Hebrew, a language with which we were able to confirm they were unfamiliar. This assessment prompt functioned as both a reliability screening device for college-aged participants, while also measuring children's ability to engage in metacognitive reflection regarding their own understanding. As such, this metacognitive assessment question was potentially informative as a screening tool, allowing us to better clarify the nature of age-related change in children's sociolinguistic awareness. Specifically, directly assessing metacognition enabled us to disentangle age-related change in the phenomenon of interest (differentiation between disability and foreignness concepts; disability spread) from variability attributable to more general metacognitive difficulty, thus shedding more nuanced light on the cognitive foundations of children's developing sociolinguistic awareness.

In sum, the overarching goal of these studies was to glean more fully articulated information about a) the cues children and college students are sensitive to in other children's speech and appearance, and b) the socially significant inferences they draw based on these cues. The methodological approach taken to assess these goals is promising as it supports multiple

comparative analyses, including across social concepts (disability and foreignness), modalities (speech and appearance), and ages (wide range of children as well as emerging adults). Ultimately, this approach may help provide an empirical basis for evaluating the conceptual basis of stigma and point to potential avenues for intervention aimed at reducing the expression of prejudice towards stigmatized others.

CHAPTER II

AN EXPLICIT APPROACH TO ASSESSING FOREIGNNESS AND DISABILITY CONCEPTS

2.1. Overview

As noted, previous research has neglected to determine the extent to which children distinguish between a speech disorder and a foreign accent when drawing generalizing inferences about speakers with a speech disorder or accent. In fact, at least under certain circumstances, previous research suggests that even expert listeners may have difficulty distinguishing between an accent and speech disorder (Ghio et al., 2013). Similarly, despite theoretical accounts linking children's reasoning about disability to their reasoning about other out-groups (Babik and Gardner, 2021), we are unaware of any existing empirical work directly assessing the merits of this account, nor any empirical exploration of developmental change in this regard.

To address these gaps, in study 1 we assessed college students (study 1a) and young children's (study 1b) tendency to distinguish between multimodal foreignness and disability cues, the inductive generalizations associated with such differentiation, and the relation between these inductive generalizations and expression of prejudice. As noted, in study 1, we adopted an explicit indexically-based approach to assessing these issues.

First, to assess conceptual differentiation, we simply asked participants whether speakers with a foreign accent (and images of children in foreign garb) or speech disorder (and images of children with physical disabilities) were "from America" or "from another country far away". This binary probe supported the ability to assess whether children and college students differ in their tendency to attribute one sort of perceptual variability (i.e., a foreign accent and foreign garb) but not another (i.e., a speech disorder or physical disability) to ethnolinguistic group

membership (as opposed to disability status). To further probe conceptual differentiation, we assessed the extent to which participants would indicate that children with physical disabilities, but not children in foreign garb, are "good at running". This comparison was informative of the degree to which detection of a physical disability, as opposed to foreignness, related to perceptions of reduced physical competence.

Second, to measure disability spread, we assessed the extent to which participants would indicate that speakers with a speech disorder and images depicting children with physical disabilities "needed help to do things." We then compared these responses to participants' evaluations of the extent to which speakers with a foreign accent or images depicting children in foreign garb "needed help to do things." These comparisons were informative of the extent to which participants selectively associated high dependence with disability cues (but not foreignness cues) in speech and appearance. To further assess disability spread, we compared the extent to which participants would indicate that speakers with a speech disorder were "good at running" in comparison to speakers with a foreign accent. Taken together, comparing participants attributions of physical competence and dependence across stimuli was informative of disability spread and the relation between disability spread and inductive generalizations in the case of ethnolinguistic variability.

Notably, in the framework of study 1, attributing reduced physical competence to individuals with physical disabilities which directly impact movement (i.e., amputee and wheelchair images) reflects conceptual understanding but not disability spread. This is because the term disability spread entails perceptions of disability that are unrelated to an observed impairment. In contrast, attributing reduced physical competence to speakers with a speech disorder does constitute disability spread since there is no direct relationship between a speech

disorder and physical competence. As well, in our framework, attributing enhanced dependence to individuals with physical disabilities and speakers with a speech disorder can also constitute disability spread as localized perceived impairments may be translated to global perceptions of dependence that extend beyond the limitation stemming from the perceived impairment alone.

Finally, our design also supported an assessment of peer rejection (a term we will use interchangeably with prejudice) and the relation between inductive generalization tendencies and expression of peer rejection. Specifically, we asked participants whether they were interested in "being friends" with speakers with a speech disorder or foreign accent (as well as images depicting physical disabilities and foreign garb). We also assessed whether perception of reduced physical competence, enhanced dependence (or reduced comprehensibility for speech) were predictive of participants' friendship-oriented intentions across the stimuli categories. Beyond assessing disability- and foreignness-related peer rejection, this approach provided an empirical basis to assess the relation between disability spread and the expression of peer rejection.

To this end, we assessed college students' (n=180) and children's (n = 163) evaluations of three speech categories: 1) neurotypical North American English-accented speech (henceforth "L1"), 2) American English-accented speech produced by children with autism spectrum disorder (henceforth "ASD") and 3) neurotypical Spanish-accented English (henceforth "L2"), and four illustration categories: children who appeared 1) wheelchair-bound with signs of contracture in the wrist and torticollis in the neck, 2) amputees, 3) foreign in their dress, or 4) non-disabled. Participants evaluated the children represented by these speech and visual stimuli along five dimensions: 1) foreignness, 2) running ability, 3) dependence, 4) prejudice, and 5) comprehensibility (for speech).

2.2. Research Questions

In this study, we examined the following research questions:

2.2.1. Conceptual differentiation:

2.2.1.1. To what extent will participants selectively associate physical disability cues (i.e., amputation, wheelchair) but not foreignness cues (i.e., foreign garb) with reduced physical competence (running ability)?

2.2.1.2. To what extent will participants selectively associate foreignness cues (i.e., foreign accent and garb) but not disability cues (i.e., amputation, wheelchair, disordered speech) with being from another country?

2.2.2. Disability spread:

2.2.2.1. To what extent will participants selectively associate disability cues (i.e., amputation, wheelchair, disordered speech) but not foreignness cues (i.e., foreign accent and garb) with greater dependence?

2.2.2.2. To what extent will participants selectively associate disability cues in speech (in this case, speech produced by children with ASD) but not foreignness cues (i.e., L2 speech) with reduced running ability?

2.2.3. Peer rejection:

2.2.3.1. To what extent will participants express reduced interest in pursuing friendships with foreign or disabled targets (as indicated by speech and/or visual appearance) in comparison to non-foreign, non-disabled targets?

2.2.3.2. Is peer rejection most closely associated with ratings of physical competence or dependence (for visual and speech stimuli), or comprehensibility (for speech stimuli)?

2.3. Predictions

The preregistration for this study along with detailed predictions is available at https://osf.io/bqut5/. In short, we predicted that college students and children alike would selectively differentiate between disability and foreignness cues vis-à-vis foreignness judgments and would selectively attribute greater dependence and reduced competence to disability stimuli. In contrast, we predicted that college students and children would differ in their friendship ratings. Namely, we predicted that college students would show no preference towards any of the image categories but would express prejudice against ASD speakers (but not L2 speakers). In contrast, we predicted that children would prefer neurotypical non-foreign images over the other three image categories (foreign, wheelchair, amputee) and would also prefer L1 speakers over ASD and L2 speakers (thus expressing prejudice towards disability and foreignness stimuli in the case of visual and speech stimuli).

2.4. Method

2.4.1. Participants

For study 1a, we recruited 180 participants enrolled in psychology courses at the University of Oregon, who received course credit in exchange for their participation.
For study 1b, an *a priori* power analysis (based on effect sizes that are typical in the literature on children's social categorization) utilizing G*Power (Faul et al., 2009) indicated that we would need 54 children to detect a medium effect size (with 80% power) pertaining to foreignness judgments based on accent (e.g., Kinzler and Dejesus, 2013). Furthermore, our power analysis revealed that we would need 70 children to detect a medium-sized relationship between age and variability in children's conceptual representation of social categories (e.g., nationality; Hussak and Cimpian, 2019).

However, due to the exploratory nature of this study and our interest in potential moderators, we increased the sample to 182 participants, who were recruited within the Living Laboratory® at the Oregon Museum of Science and Industry (OMSI) in Portland, Oregon. Data from 18 children were discarded due to repeated failure to respond to prompts, loss of interest in participating, or parental intervention, resulting in a sample of 164 participants. Since we recruited only one child under three, we removed their data from our analysis resulting in a sample of 163. Since the caregivers of 15 participants reported impossible ages for their children, their data were omitted from all age-related analyses. The final sample was thus 163 for all non-age-related analyses (89 girls, 74 boys) and 148 (80 girls, 68 boys, M_{age} = 70 months, SD = 19 months, age range = 38:119 months) for all age-related analyses.

2.4.2. Materials

2.4.2.1.Speech Stimuli

Speech samples for this study were produced by children aged 7-11 and were obtained from two primary sources:

First, speech disorder and neurotypical speech samples were obtained from Redford and colleagues (2018). The speech disorder samples were recordings of children diagnosed with autism spectrum disorder (ASD) with normal cognition whose speech featured atypical prosody. The neurotypical samples were recordings of children who participated in a longitudinal study on the typical acquisition of prosody in school-aged children. Structured spontaneous speech samples were obtained using a storytelling task based on books that depicted the activities of a frog and/or a boy and a dog (the frog story books by Mercer Mayer). For neurotypical speech samples, children were asked to select one of four picture books to narrate. They were then given an opportunity to familiarize themselves with the book they selected and told their story

two times to a parent or caregiver. Stories were told twice to allow children to plan their story and practice the language they would use to tell the story. Clips from the second storytelling event were therefore used for the present study. In contrast, since children with ASD resisted telling the same story twice, the elicitation procedure was modified. Namely, children chose one of the four books to narrate and then developed a story in response to questions and prompts from the experimenter while looking through their book of choice. The child was then instructed to "tell their story" to their caregiver. This modification provided children with ASD with a similar opportunity to practice the language they would use in advance without requiring them to tell the story twice (see Redford et al., 2018 for further details).

Second, Spanish-accented speech samples were obtained from Llanes (2012) and were comprised of Catalan/Spanish bilingual children who were learning English as a second language. Structured spontaneous speech samples were obtained through a picture-elicited narrative task (Heaton, 1966) that had been successfully used in previous work with children (Muñoz, 2006; Tavakoli and Foster, 2011). In this task, participants were asked to explain a story that consisted of six pictures (e.g., children packing a picnic only to discover that their dog climbed into the picnic basket; children playing with a ball that falls into a pit and their efforts to retrieve the ball). Participants were given one minute to have a look at the story and plan their utterances.

Thus, while the elicitation procedures differed across the three speech categories, they were all obtained in a storytelling context in which they were given an opportunity to rehearse their speech. Furthermore, the details of the stories told were similar in that they involved children's activities with animals such as dogs. For each speaker, three uninterrupted speech clips lasting between 8 and 10 seconds were selected by the author (NW). The final set of

stimuli consisted of 6 speakers (3 male, 3 female) in each of the three speech categories (Neurotypical-native English accent, Neurotypical-Spanish accent, ASD). In total, the speech stimuli thus consisted of 18 speakers and 54 speech clips.

2.4.2.2.Illustration Stimuli

A commissioned illustrator (http://www.springerdesign.biz) created visualizations that varied along two dimensions: foreignness and disability status. For foreignness, illustrations depicted: 1) children with mainstream North American clothing (e.g., shorts, t-shirt and sneakers) or 2) children with clothing intended to evoke (stereotypical) notions of Spanishness (e.g., traditional Spanish flamenco dress or a Matador's traje de luces). For disability status, illustrations depicted two examples of physical disability: 1) children with an amputated leg, and 2) children in wheelchairs with signs of contracture in the wrist and torticollis in the neck. The final stimuli set consisted of 6 children (3 male, 3 female) in each of the four visual stimuli categories (typical, amputee, wheelchair, and foreign). Notably, to optimize control, the same individuals (i.e., facial features, skin tone) were depicted across the four visualization categories. All illustrations can be found in the supplementary materials (section 1).

2.4.3. Measures

Participants were presented with audio stimuli and then visual stimuli (or the opposite order), and for each stimulus item were asked to respond to the following prompts:

2.4.3.1.Open impression - Participants were asked: "What stands out to you about this child? How would you describe them? Do you notice anything?".

2.4.3.2. Foreignness - Participants were asked to indicate (Forced-Choice) whether the speaker/child in the illustration/audio clip is "from America" or "another country far away".

2.4.3.3. Physical Competence – Participants were asked to indicate on a 3-point Likert scale ("Not really", "A little bit", "Very much") whether the speaker/child in the image/audio clip is "good at running".

2.4.3.4. Dependence - Participants were asked to indicate on a 3-point Likert scale ("Not really", "A little bit", "Very much") whether the speaker/child in the illustration/audio clip "needs a lot of help to do things".

2.4.3.5. Comprehensibility (speech samples only) - Participants were asked to indicate on a 3-point Likert scale ("Not really", "A little bit", "Very much") whether it is"easy to understand what they are saying".

2.4.3.6. Peer rejection - Participants were asked to indicate on a 3-point Likert scale ("Not really", "A little bit", "Very much") whether they "would like to be friends" with the speaker/child in the illustration/audio clip.

2.4.3.7. Reliability check - after completing the task, participants listened to a brief speech segment in Hebrew (a language they were very unlikely to be familiar with) and we gathered comprehensibility scores by asking whether it is "easy to understand what they are saying?" regarding the Hebrew audio clip (3-point Likert scale: "Not Really", "A Little Bit", "Very Much").

2.4.4. Procedure

All procedures followed the guidelines of the ethical advisory board at the University of Oregon. For study 1a, upon granting their consent, college students provided demographic information regarding their parent's levels of education as well as their age, gender, exposure to disability and mono/multilingualism. For study 1b, this information was provided by the caregivers. In both studies, the experiment was implemented/hosted on Gorilla (Gorilla.sc,

Anwyl-Irvine et al., 2020). For study 1a, college students participated online without experimenter moderation. In contrast, for study 1b, testing took place on a one-to-one basis with a researcher on the floor at the Oregon Museum of Science and Industry (OMSI), following Living Lab guidelines (National Living Lab Initiative | National Living Laboratory).

The order of stimuli presentation (audio/visual) was counter balanced. Following an open-impression prompt for each stimulus, the remaining measures/prompts were presented in randomized order. The study took approximately 10 minutes to complete and participants were compensated with research credits (in study 1a) or stickers (in study 1b) for completing the study.

2.4.5. Analysis

Our analysis plan was preregistered at https://osf.io/bqut5/. Mixed-effects analyses were performed in the R statistical language (Version 4.1.3; R Core Team. 2020) using the lme4 package (Bates et al., 2015), and *p*-values of the models were calculated using the lmerTest package (Kuznetsova et al., 2017). In these models, categorical variables were contrast coded such that the intercept of a model represents the mean of all data points in the data set. All continuous predictors were standardized and mean-centered to facilitate the interpretation of intercepts. *Post-hoc* analyses were conducted through pairwise comparisons using the emmeans package (Lenth, 2022).

We modeled all forced-choice dependent variables using mixed-effects logistic regression and all continuous dependent variables with mixed-effects linear regression. To assess within category heterogeneity, we added a nested random intercept term for speaker in each model (1 | audio type/speaker). As well, since each participant listened to all three speech conditions, we added a random intercept for order (1 | trial_number) to account for potential

order effects on each dependent variable. We then ran likelihood ratio tests comparing models with/without these random effects for each dependent variable and selected the model that provided the best fit for the data. We followed the same procedure for analysis of visual stimuli. We set neurotypical American English (L1) as the reference group for all audio-related regression analyses and non-disabled children in mainstream American clothing as the reference group for all visual-related regression analyses. Alpha was set at .05 when examining the relationship between variables but we discuss the overall support for a given hypothesis based on the pattern of findings observed across analyses. Finally, we used the Tukey method to correct for multiple comparisons in our *post hoc* analyses.

CHAPTER III

STUDY 1a (COLLEGE-AGED SAMPLE) RESULTS AND DISCUSSION

In this study, we assessed college students' (n=180) evaluations of three speech categories (neurotypical North American English, neurotypical Spanish English, and North American English produced by children with ASD) and four illustration categories (two with a disability: wheelchair, amputee, and two without disability: North-American, foreign) along five dimensions (national origin, comprehensibility for speech, running ability, perceived need-of-help, and interest in friendship). The overarching goal of our analysis was to 1) assess college-students tendency to distinguish between multimodal foreignness and disability cues, 2) compare college-students inductive generalizations based on foreignness and disability cues, and 3) assess how these generalizing tendencies relate to expression of stigma and peer rejection.

As noted above, we predicted that college students would selectively associate reduced physical competence and enhanced dependence with images depicting physical disabilities as well as ASD speakers in a similar manner to which they would associate L2 speech and foreign images with "being from another country far away". We also predicted that college students would show no preference for either of the images but would show prejudice against ASD speakers.

Descriptive statistics for parental education and participant's age, gender, exposure to disability, multilingualism, race, and ethnicity can be found in the supplementary materials (section 2).

3.1. Data Cleaning and Reliability Check Assessment

Of the 180 participants who were assessed, 6 participants (3.3%) responded in a singular manner to all Likert-scale audio prompts and 2 (1.1%) participants responded in a singular

manner to all Likert-scale visual prompts. Furthermore, seven out of the eight participants who responded in a singular manner also indicated that Hebrew was "a little" easy for them to understand. Taken together, these participants clearly did not approach the task with sufficient attention and their data were removed from further analysis.

Of the remaining 172 participants, 153 (88%) indicated that it was not easy for them to understand Hebrew. In contrast, 6 (3.4%) indicated that Hebrew was very easy to understand and 13 (7.4%) indicated that Hebrew was a little easy to understand and the response of two participants were not recorded. Preliminary analysis using reliability check as a moderator indicated that there were no meaningful differences between speech conditions (and visual conditions) on any of our outcome measures for participants who did not report difficulty understanding Hebrew. Furthermore, since the study was self-administered online without mediation, the reliability-check supported attempts to identify unreliable responses. Therefore, we decided to drop the participants who did not report difficulty understanding Hebrew along with the participant for whom this response was unavailable, resulting in a final sample size of 152 (103 Female, 45 Male, 1 gender fluid, two transgender, 1 other). Analysis of the aggregate sample, revealing similar patterns across research questions, can be found in the supplementary materials (section 5).

3.2. Open-Ended Responses

Unfortunately, since the experiment was run online without active mediation, only six participants provided intelligible open-ended responses regarding the stimuli. Of these six participants, two provided responses that were informative of perceptual attunement to differences across the three speech categories. Specifically, regarding ASD speakers, these participants noted "the slowness and youngness of the voice – the speaker is not yet confident in

speaking" and that "they sound young and unclear and really breathy". In contrast, for L2 speakers, participants commented on the accent. Finally, for L1 speakers, participants noted that speakers were "well spoken" and sounded "like a typical young boy".

For the images, all participants who provided comments regarding the wheelchair and amputee images pointed to the disability (e.g., "This person is in a wheelchair."; "They have an amputated leg. I would describe them as disabled."). Furthermore, regarding the wheelchair image, one participant indicated that "they look like they have cerebral palsy or something similar". In contrast, foreign images were described as "unique", "eloquent," and "cool" but terms like foreign were not noted. Comments regarding the typical images were generally neutral (e.g., "nothing", "a boy"), though one noted that they look "athletic".

Thus, the open-ended responses provided some interesting insights regarding college students' spontaneous reactions to the visual and speech stimuli. However, due to the low number of responses, generalized conclusions are not warranted.

3.3.Research questions

For each of the primary research questions considered below (conceptual differentiation, disability spread, prejudice), relevant descriptive statistics, the results of mixed-effects models and *post-hoc* pairwise comparisons, and a visualization are provided, followed by narrative explication of the findings.

3.3.1. Conceptual differentiation:

3.3.1.1.To what extent will participants selectively associate physical disability cues (i.e., amputation, wheelchair) but not foreignness cues (i.e., foreign garb) with reduced physical competence (running ability)?

Table 1. Descriptive Statistics for physical competence (running) scores on 3-point Likert scale.

Physical Competence Scores (Visual)		
Category	Mean	SD
Typical	2.28	0.50
Amputee	1.09	0.29
Wheelchair	1.09	0.38
Foreign	2.03	0.49

Table 2. Coefficient estimates from the mixed effects model predicting physical competence scores from image category (Note: North American-presenting images without disability are the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	2.28	0.03	65.39	< 0.0001
Amputee	-1.18	0.05	-25.30	< 0.0001
WC	-1.19	0.05	-25.44	< 0.0001
Foreign	-0.25	0.05	-5.34	< 0.0001

Table 3. Post hoc pairwise comparisons assessing differences in physical competence scores between image categories

contrast	*estimate*	*SE*	*t*	*p*
Typical - Amputee	*1.1842*	*0.0467*	*25.3040*	*<0.0001*
Typical - WC	*1.1907*	*0.0467*	*25.4446*	*<0.0001*
Typical - Foreign	*0.2500*	*0.0467*	*5.3419*	*<0.0001*
Amputee - WC	*0.0065*	*0.0467*	*0.1405*	*0.99*
Amputee - Foreign	*-0.9342*	*0.0467*	*-19.962*	*<0.0001*
WC - Foreign	*-0.9407*	*0.0467*	*-20.1026*	*<0.0001*

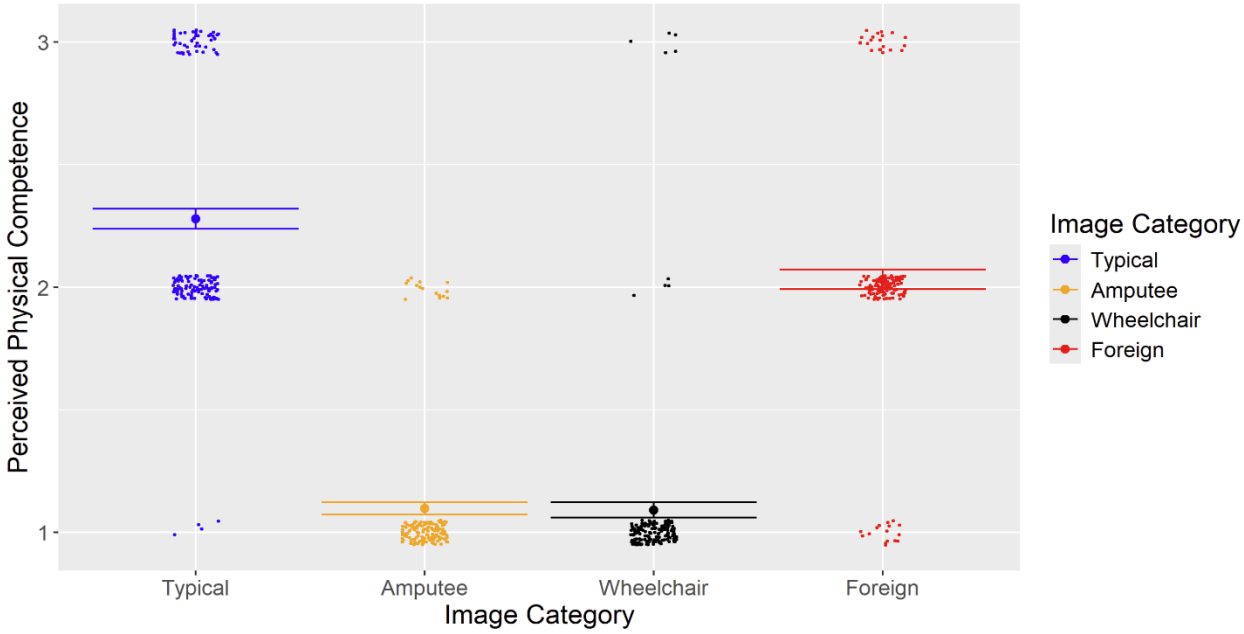

Figure 1. **Perceived Physical Competence scores across the four image categories. "Typical" refers to North American presenting images without disability, "Amputee" refers to North American presenting images with amputation, "Wheelchair" refers to North American presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Error bars represent standard errors around the mean.**

For the visual stimuli, participants clearly provided significantly higher running scores for non-disabled North American-presenting ("Typical") and foreign-presenting images in comparison to amputee and wheelchair images (see Tables 1-3 and Figure 1). In addition, foreign-presenting images received significantly lower running scores than typical images but significantly higher running scores than the two image categories depicting disabilities (see Table 3). There was no significant difference between the amputee and WC images with respect to running scores (see Table 3).

While the foreign illustrations received significantly lower running scores than typical images, they were not below the median score of 2, meaning that foreign illustrations were rated as "a little bit good at running" (see Table 1 and Figure 1). Thus, participants clearly selectively

attributed low physical competence to the disability images, as predicted. This finding helps to validate that the visual stimuli indeed adequately depicted physical disability for our college student participants.

3.3.1.2.To what extent will college students selectively associate foreignness cues (i.e., foreign accent and garb) but not disability cues (i.e., amputation, wheelchair) with being from another country?

Table 4. Descriptive Statistics and binomial tests for foreignness judgments (binary choice: "America" or "Another country far away")

Foreignness Judgments (Visual)				Foreignness Judgments (Speech)		
Category	% "far away" selection	Two-tailed Binomial significance test		Category	% "far away" selection	Two-tailed Binomial significance test
Typical	18.4%	$p < 0.0001$		L1	3.2%	$p < 0.0001$
Amputee	18.4%	$p < 0.0001$		L2	83.5%	$p < 0.0001$
Wheelchair	12.5%	$p < 0.0001$		ASD	15.7%	$p < 0.0001$
Foreign	75%	$p < 0.0001$				

Table 5. Coefficient estimates from the logistic mixed effects model predicting foreignness judgments from image category (Note: North American-presenting images without disability are the reference group).

	Estimate	*SE*	*z*	*p*
Intercept	-8.06	0.868	-9.281	< 0.0001
Amputee	0.0000005	0.628	0.000	0.99
WC	-2.134	0.751	-2.841	0.00449
Foreign	15.025	1.373	10.942	< 0.0001

Table 6. Post hoc pairwise comparisons assessing differences in foreignness judgments between image categories

contrast	*estimate*	*SE*	*z*	*p*
Typical - Amputee	*-0.0000005*	*0.628646*	*-0.0000008*	*0.99*
Typical - WC	*2.1347674*	*0.751287*	*2.8414772*	*0.0269*
Typical - Foreign	*-15.02539*	*1.373170*	*-10.94212*	*< 0.0001*
Amputee - WC	*2.134767*	*0.751314*	*2.8413780*	*0.02695*
Amputee - Foreign	*-15.0253*	*1.373150*	*-10.94227*	*< 0.0001*
WC - Foreign	*-17.1601*	*1.572704*	*-10.91124*	*< 0.0001*

Table 7. Coefficient estimates from the logistic mixed effects model predicting foreignness judgments from speech category (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	SE	*z*	*p*
Intercept	-12.88	1.336	-9.637	< 0.0001
L2	19.534	1.824	10.709	< 0.0001
ASD	6.166	0.9077	6.793	< 0.0001

Table 8. Post hoc pairwise comparisons assessing differences in foreignness judgments between speech categories

contrast	*estimate*	*SE*	*z*	*p*
L1 - L2	*-19.534*	*1.8241*	*-10.708*	*< 0.0001*
L1 - ASD	*-6.1659*	*0.9077*	*-6.792*	*< 0.0001*
L2 - ASD	*13.368*	*1.2382*	*10.796*	*< 0.0001*

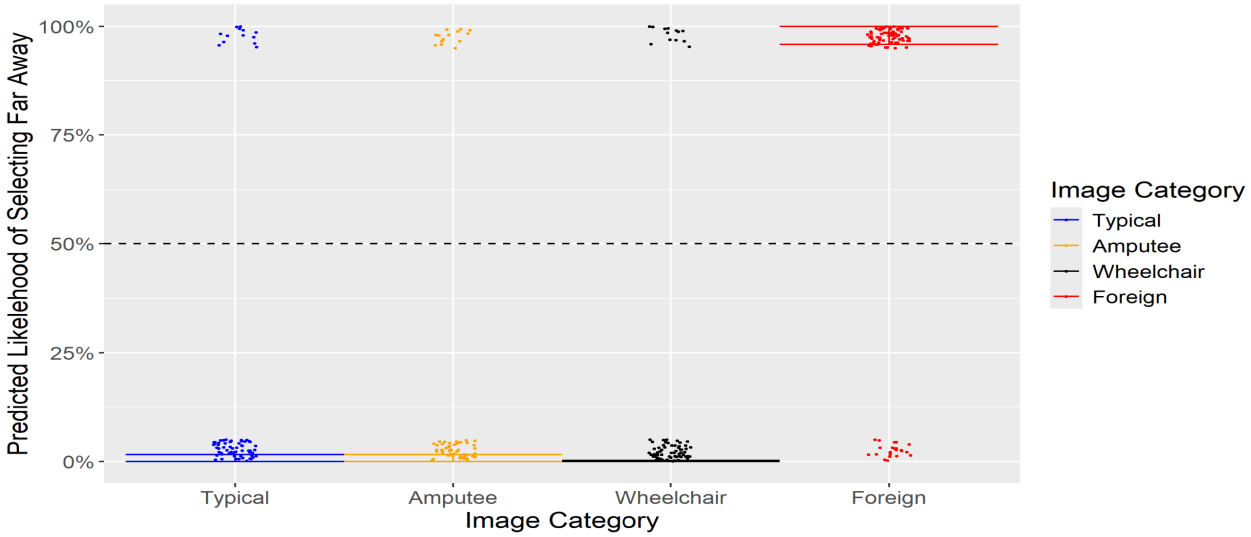

Figure 2. **Predicted likelihood of selecting "far away". "Typical" refers to North American-presenting images without disability, "Amputee" refers to North American-presenting images with amputation, "WC" refers to North American-presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Error bars represent 95% confidence intervals computed from the mixed-effects regression model.**

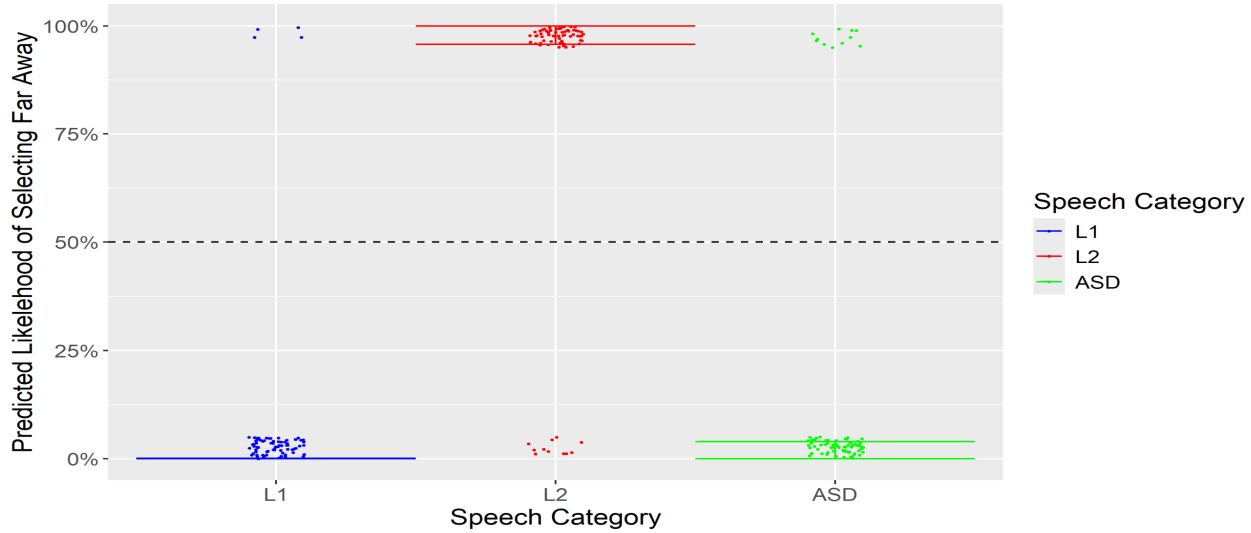

Figure 3. **Predicted likelihood of selecting "far away". L1 refers to neurotypical North American English speakers, L2 refers to neurotypical Spanish speakers learning English as a second language, ASD refers to North American English speakers with autism spectrum disorder. Error bars represent 95% confidence intervals computed from the mixed-effects regression model.**

For the visual stimuli, college students clearly reserved "far away" ratings for foreign-presenting images in comparison to all three of the other image categories, as predicted (see Tables 4-6). Unexpectedly, they were also more likely to say that WC images were from America in comparison to both the typical and amputee images (see Table 6), though they overwhelmingly indicated that typical and amputee were also from America (see Table 4).

For speech stimuli, college students clearly identified L2 speakers as more likely to be from "far away" than L1 speakers and ASD speakers, as predicted. (see Tables 4, 7 and 8). At the same time, unexpectedly, they were also significantly more likely to identify L1 speakers as being "from America" in comparison to ASD speakers (see Table 8). Despite the significant difference between L1 and ASD foreignness judgments, participants still indicated that ASD speakers were "from America" at levels significantly higher than chance, as predicted (see Table 4).

In sum, college students clearly selectively associated foreignness cues (i.e., foreign accent and garb) but not disability cues (i.e., amputation, wheelchair) with being from another country. This finding provides additional validation for the foreign stimuli (both visual and auditory) utilized in the study as discriminable representations of foreignness to college students. Participants' responses also clearly indicated that they judged ASD speakers to be from America, which points to college students recognizing that some forms of atypical speech, such as speech indicative of ASD, are distinguishable from atypical speech that is indicative of non-native language background. All in all, participants' foreignness judgments displayed clear differentiation between physical disability cues and foreignness cues, as predicted.

3.3.1.3. Conceptual differentiation Summary

Participants selectively associated illustrations featuring visible disability cues (i.e., amputation, wheelchair) with reduced physical competence (running ability) in comparison to illustrations featuring non-disabled North American and foreign-presenting children. Participants also selectively associated illustrations of children featuring foreign garb as well as speakers with a foreign accent with being from another country "far away". Taken together, these findings provided clear evidence of conceptual differentiation in college students' attunement to disability and foreignness cues in both appearance and speech, as predicted.

3.3.2. Disability spread:

3.3.2.1. To what extent will college students selectively associate disability cues (i.e., disordered speech, amputation, wheelchair) but not foreignness cues (i.e., foreign accent and garb) with greater dependence?

Table 9. Descriptive Statistics for dependence (needing-help) scores on 3-point Likert Scale

Dependence Scores (Visual)				Dependence Scores (Speech)		
Category	Mean	SD		Category	Mean	SD
Typical	1.43	0.53		L1	1.80	0.69
Amputee	2.13	0.59		L2	1.68	0.61
Wheelchair	2.58	0.57		ASD	2.13	0.70
Foreign	1.34	0.50				

Table 10. Coefficient estimates from the mixed effects model predicting perceived dependence scores from image category (Note: North American-presenting images without disability are the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	1.43	0.04	31.99	< .001
Amputee	0.70	0.06	11.57	< .001
WC	1.15	0.06	18.92	< .001
Foreign	-0.09	0.06	-1.51	.131

Table 11. Post hoc pairwise comparisons assessing differences in dependence scores between image categories

contrast	*estimate*	*SE*	*t*	*p*
Typical - Amputee	*-0.7039*	*0.06086*	*-11.565*	*< 0.0001*
Typical - WC	*-1.1513*	*0.06086*	*-18.915*	*< 0.0001*
Typical - Foreign	*0.0921*	*0.06086*	*1.5132*	*0.7854*
Amputee - WC	*-0.4473*	*0.06086*	*-7.350*	*< 0.0001*
Amputee - Foreign	*0.796*	*0.0608*	*13.079*	*< 0.0001*
WC - Foreign	*1.2434*	*0.0608*	*20.429*	*< 0.0001*

Table 12. Coefficient estimates from the mixed effects model predicting perceived dependence scores from speech category (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	1.80	0.05	33.14	< .001
L2	-0.12	0.07	-1.69	.092
ASD	0.33	0.07	4.76	< .0001

Table 13. Post hoc pairwise comparisons assessing differences in dependence scores between speech categories

contrast	*estimate*	*SE*	*t*	*p*
L1 - L2	*0.11842*	*0.06912*	*1.7131*	*0.2631*
L1 - ASD	*-0.3289*	*0.06912*	*-4.7588*	*<0.0001*
L2 - ASD	*-0.4473*	*0.06912*	*-6.4719*	*<0.0001*

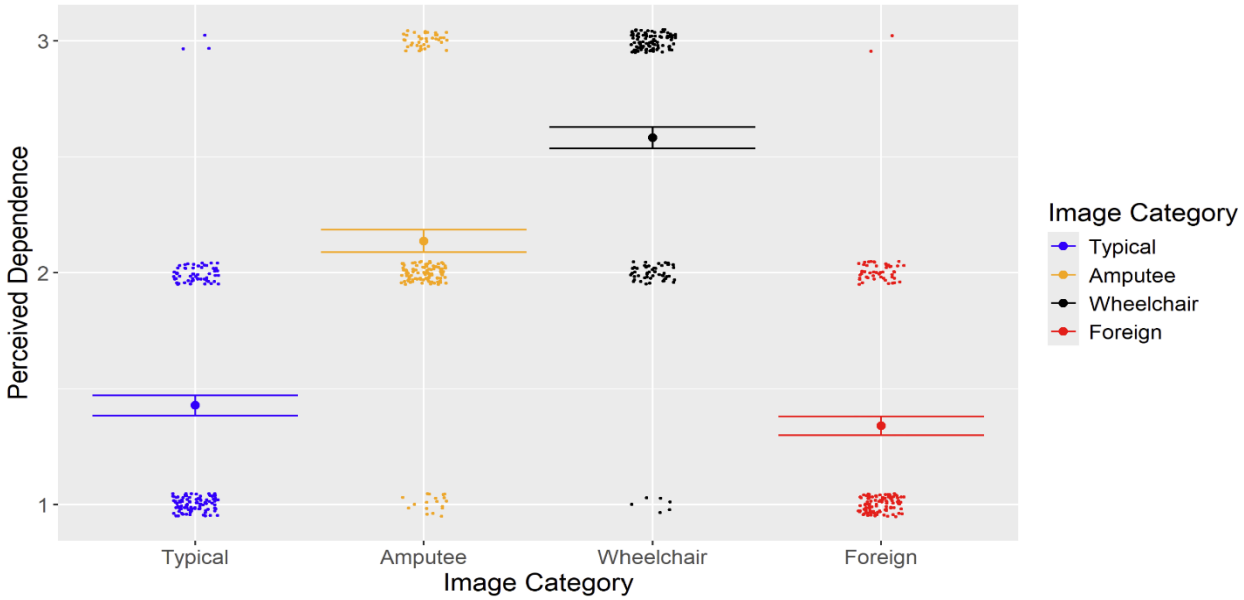

Figure 4. **Perceived dependence scores across the four image categories. "Typical" refers to North American presenting images without disability, "Amputee" refers to North American presenting images with amputation, "Wheelchair" refers to North American presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Error bars represent standard errors around the mean.**

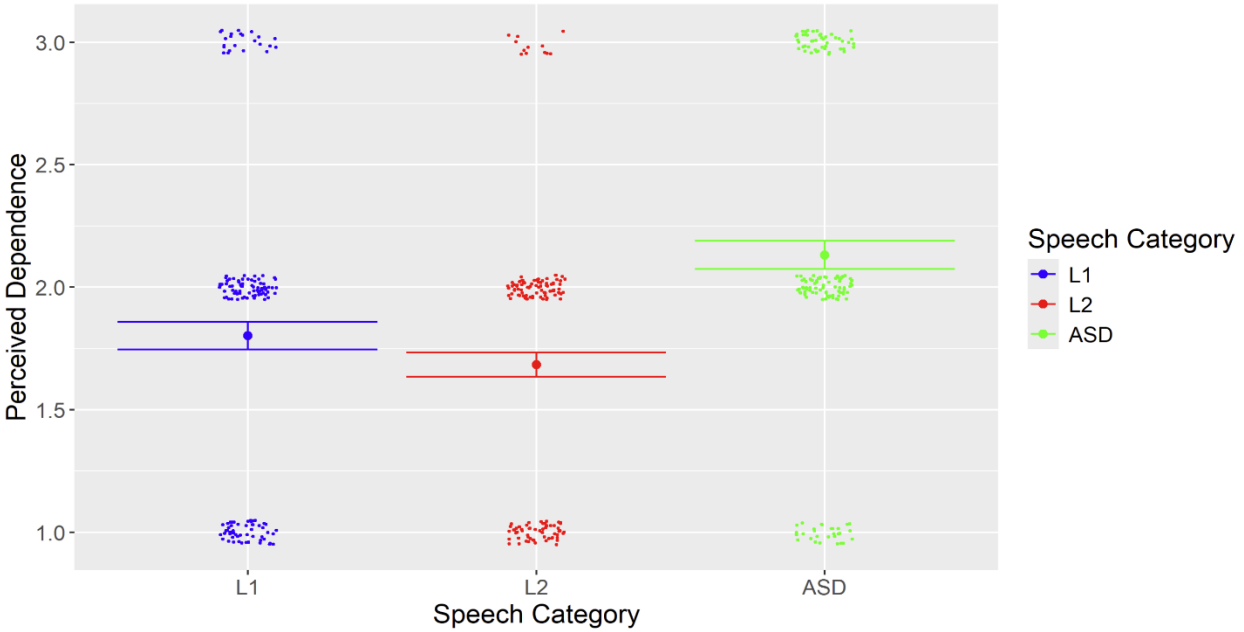

Figure 5. **Perceived dependence scores across the three speech categories. L1 refers to neurotypical North American English speakers, L2 refers to neurotypical Spanish speakers learning English as a second language, ASD refers to North American English speakers with autism spectrum disorder. Error bars represent standard errors around the mean.**

For the visual stimuli, participants clearly provided significantly lower dependence scores for typical North American-presenting and foreign-presenting images in comparison to amputee and wheelchair images (see Tables 9-11 and Figure 4). Furthermore, while the difference between foreign and typical images was not significant, WC images received significantly higher dependence scores in comparison to the amputee images (Table 11 and Figure 4). The fact that college students did not differ in their judgments of physical competence between the amputee and WC images but did provide higher dependence scores for the WC images suggests an attunement to the features depicted in the WC images that elicited a greater sense of dependence than the isolated amputee disability indicator.

For the speech stimuli, participants provided higher dependence scores for ASD speakers in comparison to both L1 and L2 while the difference between L1 and L2 was not significant (see Table 13 and Figure 5). Thus, there was clear evidence of disability spread in college students' judgments of ASD speakers, as predicted.

3.3.2.2.To what extent will college students selectively associate reduced running ability with ASD speakers but not L2 speakers?

Table 14. Descriptive Statistics for physical competence (running) scores on 3-point Likert scale.

Physical Competence Scores (Speech)		
Category	Mean	SD
L1	2.01	0.47
L2	1.92	0.47
ASD	1.79	0.52

Table 15. Coefficient estimates from the mixed effects model predicting physical competence scores from speech category (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	SE	*t*	*p*
Intercept	2.01	0.04	50.80	< .0001
L2	-0.09	0.05	-1.84	0.067
ASD	-0.22	0.05	-4.53	< .0001

Table 16. Post hoc pairwise comparisons assessing differences in physical competence scores between speech categories

contrast	*estimate*	*SE*	*t*	*p*
L1 - L2	*0.0855*	*0.0486*	*1.7573*	*0.2396*
L1 - ASD	*0.217*	*0.0486*	*4.4609*	*0.00003*
L2 - ASD	*0.131*	*0.0486*	*2.7035*	*0.0217*

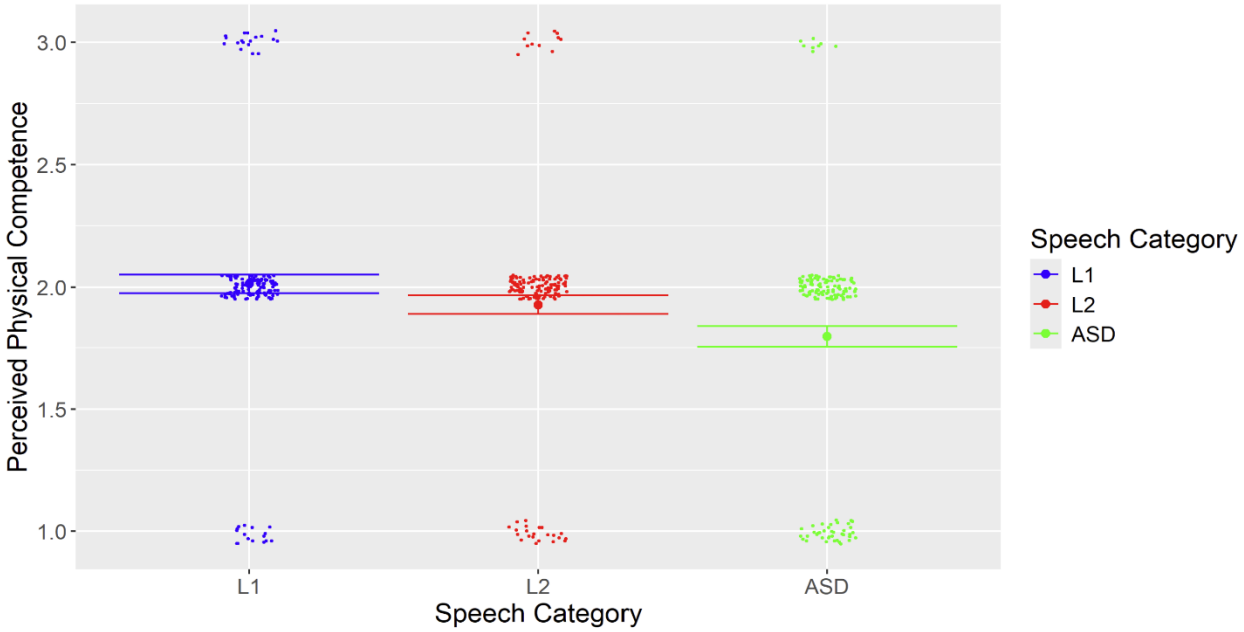

Figure 6. **Perceived physical competence scores across the three speech categories. L1 refers to neurotypical North American English speakers, L2 refers to neurotypical Spanish speakers learning English as a second language, ASD refers to North American English speakers with autism spectrum disorder. Error bars represent standard errors around the mean.**

ASD speakers received significantly lower running scores in comparison to both L1 and L2 speech categories (Table 16 and Figure 6). In contrast, there was no significant difference between running scores provided for L1 and L2 speech categories (see Table 16 and Figure 6). Thus, college students clearly associated ASD speakers with reduced physical competence, consistent with disability spread, as predicted.

3.3.2.3.Disability Spread Summary

In sum, there was clear evidence of disability spread in college students' responses to the speech and visual stimuli alike, as predicted. Namely, participants associated images depicting physical disabilities with greater dependence in comparison to the typical and foreign images. They also associated autistic speakers with greater dependence and reduced physical competence in comparison to speakers in the other two speech categories. Thus, college-students evaluated ASD speakers in a similar manner to which they evaluated images of children with physical disabilities. This raises the possibility that participants evoked a latent disability concept to account for perceived variability in the speech produced by children with ASD.

3.3.3. Peer Rejection

3.3.3.1. To what extent will participants express reduced interest in pursuing friendships with foreign or disabled targets (as indicated by speech and/or visual appearance) in comparison to non-foreign, non-disabled targets?

Table 17. Descriptive Statistics for friendship scores on 3-point Likert Scale

Friendship Scores (Visual)				Friendship Scores (Speech)		
Category	Mean	SD		Category	Mean	SD
Typical	2.10	0.64		L1	1.90	0.74
Amputee	2.25	0.53		L2	1.97	0.64
Wheelchair	2.20	0.60		ASD	1.77	0.71
Foreign	2.13	0.61				

Table 18. Coefficient estimates from the mixed effects model predicting friendship scores from image category (Note: North American-presenting images without disability are the reference group).

	Estimate	*SE*	*t*	*p*
Intercept	2.11	0.05	43.27	< .001
Amputee	0.15	0.05	3.11	.002
Wheelchair	0.10	0.05	2.03	.043
Foreign	0.03	0.05	0.54	.588

Table 19. Post hoc pairwise comparisons assessing differences in friendship scores between image categories

contrast	*estimate*	*SE*	*t*	*p*
Typical - Amputee	*-0.151*	*0.0486*	*-3.113*	*0.011*
Typical - WC	*-0.098*	*0.0486*	*-2.030*	*0.257*
Typical - Foreign	*-0.026*	*0.0486*	*-0.5414*	*0.99*
Amputee - WC	*0.052*	*0.0486*	*1.08283*	*0.99*
Amputee - Foreign	*0.125*	*0.0486*	*2.57173*	*0.062*
WC - Foreign	*0.072*	*0.0486*	*1.488*	*0.823*

Table 20. Coefficient estimates from the mixed effects model predicting friendship scores from speech category (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	1.90	0.06	33.45	< .001
L2	0.06	0.06	1.06	.288
ASD	-0.13	0.06	-2.24	.026

Table 21. Post hoc pairwise comparisons assessing differences in friendship scores between speech categories

contrast	estimate	SE	t	p
L1 - L2	-0.06578	0.05864	-1.12187	0.788
L1 - ASD	0.131578	0.05864	2.2437	0.076
L2 - ASD	0.197368	0.05864	3.365	0.002

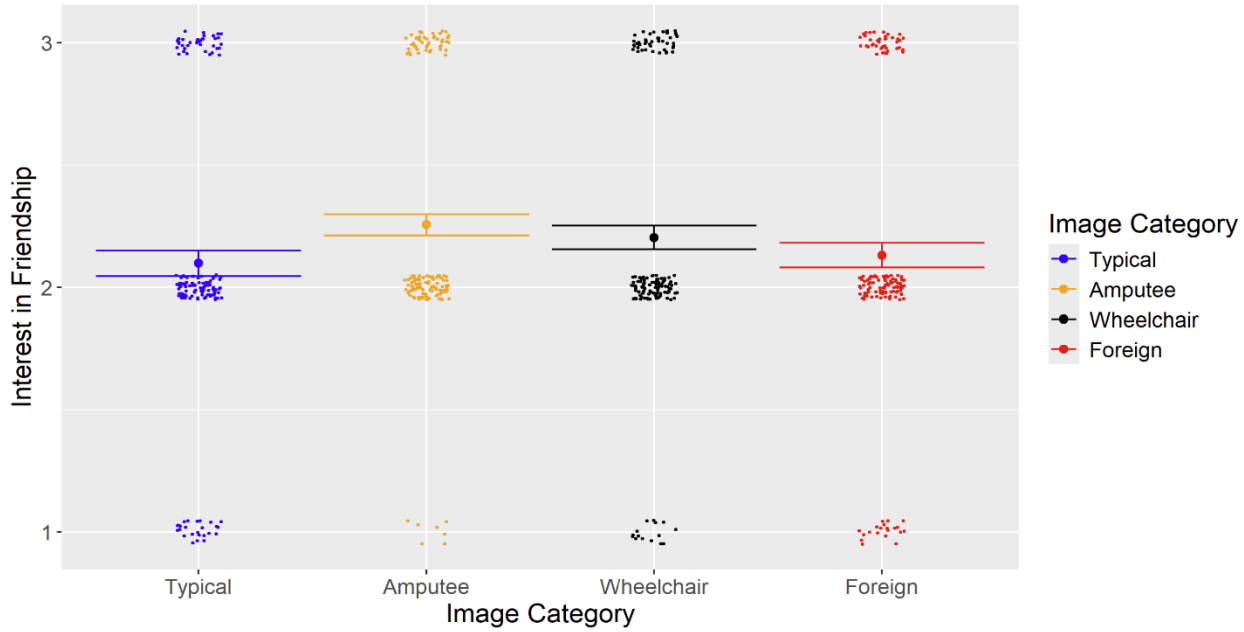

Figure 7. **Friendship scores across the four image categories. "Typical" refers to North American presenting images without disability, "Amputee" refers to North American presenting images with amputation, "Wheelchair" refers to North American presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Error bars represent standard errors around the mean.**

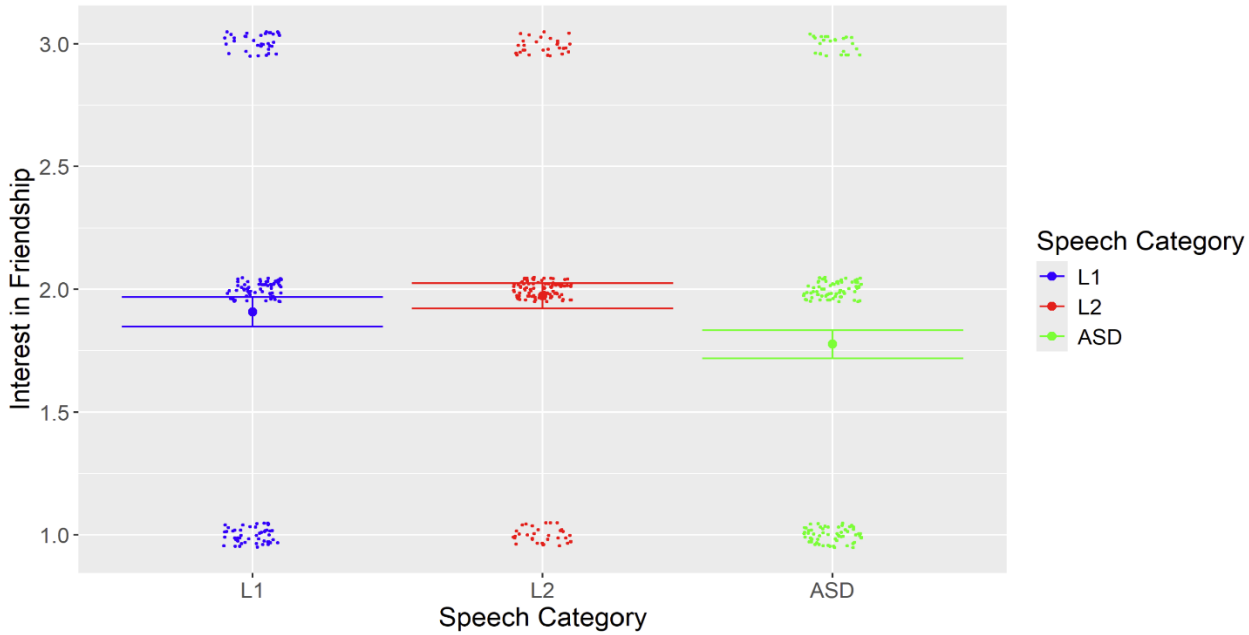

Figure 8. **Friendship scores across the three speech categories. L1 refers to neurotypical North American English speakers, L2 refers to neurotypical Spanish speakers learning English as a second language, ASD refers to North American English speakers with autism spectrum disorder. Error bars represent standard errors around the mean.**

For the visual stimuli, college students were significantly more interested in pursuing friendships with children represented by amputee images in comparison to typical images and there was a trending tendency for participants to prefer children represented by amputee images over foreign images as well (see Tables 17-19 and Figure 7). Beyond those effects, no significant differences in friendship scores were observed across the image categories (see Table 19), though there was a trending tendency for wheelchair images to receive higher friendship scores than typical images as well (see Table 18). Thus, as predicted, the college participants did not exhibit disability-related prejudice towards the children depicted by images with cues to physical disability. In fact, unexpectedly, participants showed a preference for the children represented by amputee images over the non-disabled images.

For the speech stimuli, college students were significantly more interested in pursuing friendships with L2 speakers in comparison to ASD speakers and there was a trending tendency for participants to prefer L1 speakers over ASD speakers as well, as predicted (see Tables 20-21 and Figure 8). On the other hand, there was no significant difference between L1 and L2 speakers in this regard. Thus, participants showed signs of peer rejection when evaluating speakers with ASD, but not L2, as predicted.

In sum, consistent with previous research (Redford et al., 2018), participants provided lower friendship scores for ASD speakers in comparison to the other speech categories. On the other hand, college students did not associate images depicting physical disabilities with lower friendship scores and, in fact, showed a preference towards the amputee images (and towards the wheelchair images as well, though to a lesser extent).

The tendency for college students to exhibit prejudice in the case of autistic speakers but not in the case of physical disabilities is noteworthy. While future research is needed, this

tendency is evocative of a recent proposal by Kinzler (2021) regarding contemporary expression of prejudice. On this account, individuals in developed countries may have become accustomed to curtailing their explicit expressions of prejudice towards stigmatized others. However, in the case of speech, they might still "allow" for the expression of such prejudice. Speech-related prejudice may thus operate as an invisible form of bias that continues to influence social processes.

Interestingly, Kinzler's claim was primarily focused on accents (i.e., that foreignness prejudice would be more likely to be expressed in the case of accents than visual indicators of foreignness). Thus, our data point to the possibility that these processes might be at play in the case of perceptions of disability as well. Interestingly, in contrast to Kinzler and colleagues (e.g., Kinzler et al., 2007; Kinzler et al., 2009; Kinzler and DeJesus, 2013) and other research groups (e.g., Souza., Byers-Heinlein and Poulin-Dubois, 2013; Paquette-Smith et al., 2019) we found no evidence of bias in relation to the foreign accent. Future research is needed to assess the extent to which this effect was due to the use of Spanish accents with which participants are likely familiar, to desirability bias, or to other reasons (see general discussion for further elaboration).

3.3.3.2. *Is peer rejection most closely associated with ratings of physical competence, dependence, or comprehensibility (the latter in the case of speech)?*

3.3.3.2.1. Predictors of Peer Rejection – Visual Stimuli

To assess factors contributing to the expression of peer rejection, we explored the relationship between perceived physical competence, dependence, and friendship scores for the visual stimuli. Importantly, likelihood ratio tests indicated that accounting for perceived physical competence or dependence generally did not significantly improve the model fit. This reflects the fact that these variables were generally weak predictors of friendship scores provided for the

images. However, due to our a priori interest in assessing these relationships, we provide the results of mixed effects models and *post hoc* trend analyses below.

Table 22. Coefficient estimates from the mixed effects model predicting friendship scores from the main effects of image category, physical competence scores and the interaction between these two variables (Note: North American-presenting images without disability are the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	1.81	0.18	10.04	< .001
Amputee	0.59	0.23	2.57	.010
WC	0.47	0.21	2.27	.024
Foreign	0.10	0.23	0.43	.671
Competence	0.13	0.08	1.70	.089
Amputee x Competence	-0.26	0.15	-1.73	.084
WC x Competence	-0.20	0.12	-1.62	.106
Foreign x Competence	-0.02	0.11	-0.19	.850

Table 23. Post hoc pairwise comparisons assessing the relationship between perceived physical competence and friendship scores across the image categories

Image type	Competence Trend	*SE*	lower.CL	upper.CL	*t*	*p*
Typical	0.1294	0.07614	-0.0201	0.279	1.70	0.089
Amputee	-0.1297	0.12897	-0.3831	0.1236	-1.00	0.3149
WC	-0.0706	0.09955	-0.2662	0.1248	-0.71	0.4779
Foreign	0.1094	0.0781	-0.04408	0.26306	1.4	0.1619

Table 24. Coefficient estimates from the mixed effects model predicting friendship scores from the main effects of image category, dependence scores and the interaction between these two variables (Note: North American-presenting images without disability are the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	2.23	0.11	19.53	< .001
Amputee	0.04	0.18	0.23	.822
WC	0.33	0.21	1.56	.119
Foreign	-0.20	0.15	-1.36	.175
Dependence	-0.09	0.07	-1.23	.219
Amputee x Dependence	0.08	0.10	0.84	.402
WC x Dependence	-0.05	0.10	-0.50	.620
Foreign x Dependence	0.16	0.10	1.62	.106

Table 25. Post hoc trend analysis assessing the effect of perceived dependence on friendship scores across the image categories

Image type	*Dependence Trend*	*SE*	*lower.CL*	*upper.CL*	*t*	*p*
Typical	*-0.08894*	*0.07232*	*-0.231*	*0.0531*	*-1.229*	*0.2193*
Amputee	*-0.007747*	*0.064614*	*-0.13469*	*0.1191*	*-0.119*	*0.904*
WC	*-0.13844*	*0.06798*	*-0.27201*	*-0.0048*	*-2.036*	*0.0422*
Foreign	*0.0736858*	*0.077060*	*-0.077*	*0.22508*	*0.9562*	*0.33942*

As noted, for the visual stimuli, physical competence scores were generally not predictive of friendship scores across the image conditions, though there was a trending tendency for greater physical competence to predict greater friendship scores for typical images, and to a lesser extent, for foreign images (see Tables 22 and 23). Similarly, dependence scores were generally not predictive of friendship scores, though, notably, greater dependence scores were significantly associated with reduced friendship scores provided to wheelchair images (see Table 25).

To further assess the relative contribution of each predictor (physical competence and dependence) to college students' friendship scores, we ran another model predicting friendship scores while simultaneously taking into account image category, dependence, and competence.

Table 26. Coefficient estimates from the mixed effects model predicting friendship scores from the main effects of image category, dependence, and physical competence scores (Note: North American-presenting images without disability are the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	2.12	0.14	15.30	< .001
Amputee	0.23	0.08	3.03	.003
Wheelchair	0.21	0.08	2.48	.013
Foreign	0.02	0.05	0.29	.769
Dependence	-0.06	0.04	-1.61	.108
Competence	0.03	0.05	0.55	.585

This model clarified that, even when accounting for perceived physical competence as well as perceived dependence, images depicting disabilities were significant positive predictors

of friendship scores (Table 26). Thus, to the extent that disability-related peer rejection was present (and it was generally not present), it was not tightly related to perceptions of reduced physical competence or enhanced dependence (with the possible exception of wheelchair images).

3.3.3.2.2. Predictors of Peer Rejection – Speech Stimuli

To assess factors contributing to the expression of peer rejection, we explored the relationship between perceived physical competence, dependence, comprehensibility and friendship scores for the speech stimuli.

First, we assessed the relationship between friendship and physical competence scores across the three speech conditions. Since a likelihood ratio test revealed that accounting for the interactions between speech category and physical competence scores did not significantly improve the model fit, we dropped the interaction term from the model. The null results for this interaction indicate that the relationship between physical competence and friendship scores did not significantly vary across speech conditions. Results from a mixed-effects model predicting friendship scores from the main effects of speech category and physical competence scores are shown in Table 27, and *post hoc* trend analyses illustrating the relatively uniform effect of perceived competence on friendship across the speech categories are shown in Table 28.

Table 27. Coefficient estimates from the mixed effects model predicting friendship scores from the main effects of speech category, physical competence scores and the interaction between these two variables (Note: Neurotypical North American English speech – L1 – is the reference group)

Effect	Estimate	SE	*t*	*p*
Intercept	1.34	0.13	10.04	< .0001
L2	0.09	0.058	1.56	0.12
ASD	-0.07	0.059	-1.2	0.233
Competence	0.28	0.06	4.67	<0.0001

Table 28. Post hoc trend analysis assessing the effect of perceived competence on friendship scores across the three speech categories

Speech Category	*Competence Trend*	*SE*	*lower.CL*	*upper.CL*	*t*	*p*
L1	*0.34682*	*0.09988*	*0.1504*	*0.5432*	*3.472*	*0.0005*
L2	*0.246910*	*0.10199*	*0.0463*	*0.4474*	*2.420*	*0.0159*
ASD	*0.259150*	*0.09114*	*0.0799*	*0.4383*	*2.843*	*0.0047*

This analysis revealed that physical competence judgments were significantly positively associated with friendship scores for all speech conditions (see Table 28). Furthermore, when accounting for perceived physical competence, there was no longer significant prejudice against ASD speakers (see Table 27). Thus, perceptions of reduced physical competence (which were significantly more pronounced for ASD speakers in comparison to the other speech categories – see Tables 17-18) likely played a significant role in participants' tendency to provide lower friendship scores for ASD speakers.

Next, we repeated these analyses for dependence judgements. As was the case for physical competence, a likelihood ratio test revealed that accounting for the interaction between speech category and dependence scores did not significantly improve the model fit, and we dropped the interaction term from the model. Results from a mixed-effects model predicting friendship scores from the main effects of speech category and dependence scores are shown in Table 29, and *post hoc* trend analyses illustrating the relatively uniform effect of perceived dependence on friendship across the speech categories are shown in Table 30.

Table 29. Coefficient estimates from the mixed effects model predicting friendship scores from the main effects of speech category, dependence scores and the interaction between these two variables (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	SE	*t*	*p*
Intercept	2.09	0.09	21.56	< 0.0001
L2	0.053	0.06	0.91	0.36
ASD	-0.099	0.06	-1.63	0.105
Dependence	-0.102	0.04	-2.32	0.02

Table 30. Post hoc trend analyses assessing the relationship between dependence and friendship scores across the three speech categories

Speech Category	Dependence Trend	SE	lower.CL	upper.CL	*t*	*p*
L1	-0.134536	0.06958	-0.2713	0.00229	-1.933	0.0539
L2	-0.035508	0.078151	-0.18917	0.1181	-0.4543	0.649
ASD	-0.12102	0.068030	-0.2547	0.01273	-1.7790	0.0760

This analysis revealed that there was a trending tendency for greater dependence scores to be associated with reduced friendship scores for L1 and ASD speakers (see Table 30). Furthermore, when accounting for dependence scores, there was no longer a significant prejudice effect against ASD speakers, mirroring the pattern observed with physical competence. Thus, perceptions of enhanced dependence (which were significantly higher for ASD speakers in comparison to the other speech categories – see Tables 12-13) likely played a significant role in participants' tendency to provide lower friendship scores for ASD speakers.

Finally, since prior research (e.g., Redford et al., 2018) has identified speech comprehensibility as a potentially important factor influencing people's affiliative interest in speakers with ASD, we were particularly interested in comprehensibility in the present study. Results from a mixed-effects model assessing differences in comprehensibility scores across speech conditions are shown in Table 31, and *post hoc* pairwise comparisons are shown in Table 32.

Table 31. Coefficient estimates from the mixed effects model predicting comprehensibility judgements from speech category (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	2.73	0.06	47.28	< 0.0001
L2	-1.12	0.08	-14.29	< 0.0001
ASD	-0.83	0.08	-10.69	< 0.0001

Table 32. Post hoc pairwise comparisons assessing differences in comprehensibility scores between speech categories

contrast	*estimate*	*SE*	*t*	*p*
L1 - L2	*1.1118*	*0.0782*	*14.21*	*< 0.0001*
L1 - ASD	*0.8289*	*0.0782*	*10.595*	*< 0.0001*
L2 - ASD	*-0.2828*	*0.0782*	*-3.615*	*0.001*

These analyses indicated that college students rated speakers with ASD as significantly less intelligible than L1 speakers but significantly more intelligible than L2 speakers. Notably, the fact that ASD speakers received lower comprehensibility scores than L1 speakers was consistent with the findings reported by Redford et al., (2018) in relation to these same ASD and typical North American speech (L1) stimuli.

Next, we assessed the extent to which comprehensibility scores predicted friendship scores across the three speech conditions. Results from a mixed-effects model predicting friendship scores from the main effects of speech category, comprehensibility scores, and the interaction between these two variables are shown in Table 33, and *post hoc* trend analyses are shown in Table 34.

Table 33. Coefficient estimates from the mixed effects model predicting friendship scores from the main effects of speech category, comprehensibility scores and the interaction between these two variables (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	1.15	0.25	4.57	< .001
L2	0.71	0.27	2.58	.010
ASD	0.49	0.27	1.83	.068
Comprehensibility	0.28	0.09	3.08	.002
L2 x Comprehensibility	-0.20	0.11	-1.85	.065
ASD x Comprehensibility	-0.21	0.10	-2.01	.046

Table 34. Post hoc trend analysis assessing the relationship between comprehensibility and friendship scores across the three speech categories

Speech Category	*Comprehensibility trend*	*SE*	*lower.CL*	*upper.CL*	*t*	*p*
L1	*0.27662*	*0.09013*	*0.0993*	*0.4538*	*3.0690*	*0.0023*
L2	*0.07166*	*0.0651*	*-0.05653*	*0.19987*	*1.0992*	*0.2723*
ASD	*0.06994*	*0.0563*	*-0.04076*	*0.18065*	*1.2423*	*0.2148*

Our analysis revealed that the relationship between comprehensibility and friendship scores varied significantly across the speech categories. Namely, comprehensibility scores were significantly and positively associated with friendship for L1 speech but there was no meaningful relationship between comprehensibility and friendship scores in the L2 and ASD conditions (see Table 34). Thus, reduced comprehensibility was not significantly associated with lower likeability of ASD speakers. Notably, this finding diverges from that of Redford and colleagues' (2018) finding regarding college students' judgments in relation to the ASD speech stimuli. On the other hand, when accounting for comprehensibility scores (which were significantly associated with friendship scores for L1 speakers), we observed a prosocial bias towards L2 and ASD speakers (though to a lesser extent than L2). Thus, the tendency to provide higher comprehensibility scores to L1 speakers combined with the tendency for these scores to be associated with greater friendship scores provided for L1 speakers played a significant role in the initial anti-social bias against ASD speakers that we observed (see Tables 20-21). Thus, enhanced comprehensibility associated with L1 speakers, rather than decreased comprehensibility associated with ASD speakers or L2 speakers partially accounted for the observed differences in friendship scores provided to these speakers.

To summarize the results of our analysis of speech-related predictors of friendship thus far: physical competence judgments predicted friendship scores for all three speech categories,

dependence scores trended towards predicting friendship scores for L1 and ASD; and comprehensibility judgments predicted friendship scores but only for L1.

To further assess the relative contribution of each speech-related predictor to college students' friendship scores, we ran one model predicting friendship scores from speech category, comprehensibility, dependence and competence. The results of this model are presented in Table 35.

Table 35. Coefficient estimates from the mixed effects model predicting friendship scores from the main effects of speech category, comprehensibility, physical competence and dependence scores (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	1.34	0.19	6.95	< .0001
L2	0.15	0.07	1.97	.049
ASD	0.01	0.07	0.10	.923
Comprehensibility	0.06	0.04	1.48	.139
Dependence	-0.08	0.04	-1.77	.078
Competence	0.27	0.06	4.17	0.00003

This model clarified that perceived physical competence was clearly the most significant predictor of friendship scores. Furthermore, inclusion of the physical competence variable in the model rendered the coefficient representing the difference in friendship scores between L1 and ASD to zero. Thus, perceptions of diminished physical competence were a particularly significant driver of participants' inclination to avoid friendship with ASD speakers.

3.3.3.2.3. Predictors of Peer Rejection – Summary

In sum, in college students' judgments for the visual stimuli, amputee and wheelchair images were, somewhat unexpectedly, positive predictors of friendship scores. As well, while perceived physical competence and dependence were not robust predictors of friendship scores across the image conditions, there were trending positive associations between perceived physical competence and friendship scores provided to typical images (and, to a lesser extent,

foreign images) and greater dependence scores were significantly associated with reduced friendship scores provided to the wheelchair images.

For the speech stimuli, higher physical competence scores predicted higher friendship scores for all three speech categories, in contrast to the generally null pattern observed for competence vis-à-vis the visual stimuli. On the other hand, consistent with the visual stimuli, higher dependence scores trended towards predicting lower friendship scores for L1 and ASD. Lastly, while comprehensibility predicted friendship scores only for L1, accounting for this effect resulted in a trending prosocial bias towards ASD speakers (and full-fledged prosocial bias towards L2 speakers).

The fact that greater dependence scores were associated with reduced friendship scores provided for the wheelchair images and ASD speakers is particularly informative of the relationship between spreading tendencies and the expression of prejudice. Namely, in contrast to previous theoretical accounts which have depicted disability spread as distinct from stigma (Liesener and Mills, 1999), these data highlight ways in which perceptions of dependence associated with a disability may be directly related to the expression of prejudice towards individuals with a disability (see Parker, Monteith and South, 2020). At the same time, the magnitude of such prejudice was small and further research is needed.

Beyond disability spread, the fact that ASD speakers received lower friendship scores than L2 speakers despite the latter receiving lower comprehensibility scores is also particularly noteworthy. Namely, previous accounts that highlighted the role of comprehensibility in speech-related prejudice (e.g., Perkins, 2007; Redford et al., 2018), would predict that speech categories receiving lower comprehensibility scores would also receive lower friendship scores. However, this was not the case for the present sample. In fact, when directly assessed, in contrast to

physical competence scores, comprehensibility scores were not significantly associated with friendship scores provided to autistic speakers. These findings highlight the fact that factors beyond comprehensibility alone likely undergird prejudicial reactions to individuals with a speech disorder. Nevertheless, the fact that accounting for comprehensibility resulted in a prosocial bias towards L2 and ASD speakers underscores the relevance of comprehensibility to speech-related prejudice as well.

In addition, whereas physical competence positively predicted interest in friendship for speakers across the three speech conditions, physical competence was generally not predictive of friendship for the images, and certainly not for the images depicting physical disabilities. Future research is needed to shed light on why perceived physical competence was not associated with friendship scores for the images in the way it was for speech (though see general discussion in chapter 5 for further elaboration).

3.3.4. Stimuli Validation

Beyond shedding light on college students' conceptual differentiation between foreignness and disability, disability spread, and prejudice, these findings also point to the validity of our experimental stimuli. Thus, as intended, the L2 speech and foreign garb reliably elicited perceptions of foreignness, and the images depicting physical disabilities (as well as the ASD speakers) robustly elicited perceptions of dependence and reduced physical competence. This sets the stage for our investigation of young children's evaluations of these stimuli and comparisons between young children and college students in this regard. In the next chapter, we present the results of the same study probing children's evaluations of these stimuli.

CHAPTER IV

STUDY 1b (CHILD SAMPLE) RESULTS AND DISCUSSION

In this study, we compared North American children's (ages 3-9 years) evaluations of three speech categories (neurotypical North American English (L1), neurotypical Spanish English (L2), and North American English produced by children with ASD (ASD)) and four illustration categories (two with a disability: wheelchair, amputee, and two without disability: North-American, foreign) along five dimensions (national origin, comprehensibility for speech, running ability, perceived need-of-help, and interest in friendship). The overarching goal of our analysis was to 1) assess children's tendency to distinguish between multimodal foreignness and disability cues, 2) compare children's inductive generalizations based on foreignness and disability cues, and 3) assess how these generalizing tendencies relate to expression of stigma and peer rejection.

As noted, we predicted that children would selectively associate speakers with a speech disorder (ASD) and images depicting physical disabilities with reduced competence in a similar manner to which they would associate L2 speech and foreign garb with "being from another country far away." As well, we predicted that neurotypical L1 speech samples and neurotypical North American illustrations would yield the greatest friendship scores.
Descriptive statistics for parental education and participant's age, gender, exposure to disability, multilingualism, race, and ethnicity can be found in the supplementary materials (section 2).

4.1.Data Cleaning and Reliability Check assessment

Of the 93 participants who were assessed with the reliability check, 71 (76.3%) indicated that it was not easy for them to understand Hebrew. In contrast, 17 (18.3%) indicated that Hebrew was very easy to understand and 5 (5.4%) indicated that Hebrew was a little easy to

understand. The correlation between age and these comprehensibility responses was sizable (r = .48, 95% CI: .29, .63, $p < 0.0001$). Notably, only one of the six participants who provided singular responses on more than 80% of prompts passed the reliability-check task, thus supporting the notion that this assessment would be sensitive to detecting unreliable judgment patterns.

Due to the skewed distribution in reliability-check responses and our interest in identifying participants with reliable responses, we treated the reliability-check as a categorical variable by collapsing participants who reported great ease and slight ease in understanding Hebrew. This facilitated straightforward discrimination between participants who were reliably able to report their own (lack of) comprehension from those who were not.

Preliminary analysis using the reliability check as a categorical moderator indicated that there were no meaningful differences between speech conditions (and visual conditions) on any of our outcome measures for participants who failed the reliability check. Therefore, we decided to drop the participants who failed to report difficulty understanding Hebrew along with the participants for whom these data were unavailable, resulting in a sample size of 71. Of those who passed the reliability check, the age of 7 children was inadvertently not recorded. Thus, the final sample size was 71 for all age-related analyses and 64 and for all age-related analyses. Due to the significant correlation between age and reliability check responses, removing participants who failed the reliability check generally resulted in a significant reduction in the occurrence of age-related change. Thus, we removed age as a predictor from all models in which likelihood ratio tests comparing a model with and without age indicated that accounting for age did not significantly improve the model fit. Analysis of the entire sample that involved the more robust age-related changes can be found in the supplementary materials (section 6).

4.2.Open-ended responses

To recall, for each stimuli item, participants were asked: “What stands out to you about this child? How would you describe them? Do you notice anything?”. Descriptive frequency data regarding these responses can be found in the supplementary materials (section 3). For speech, whereas some children spontaneously evoked “difference” to describe speakers with ASD (e.g., “they speak differently”; age 5.3), other responses were more deficit-oriented (e.g., “stumbled over words”; age not recorded). As well, whereas references to accents were pervasive in response to L2 speech, these were rare in the case of ASD speech and several (older) children explicitly indicated that they did not attribute the anomalies they detected in ASD speech to an accent (e.g., “it sounds like they speak my language but a little”; age 7.3). Notably, while generalized negative comments were observed in the case of speakers with ASD (e.g., “a little bit crazy”; age 4.6), these were apparent for L2 speech as well (e.g., “silly”; age 4.5, “weird”; age 8.8, and “funny talking”; age 5.3). In contrast, two older children (ages 9.6, 9.8) surmised that L1 speakers would be “good readers.”

For the visual stimuli depicting physical disabilities, wheelchairs and crutches were frequently commented on. However, while many (particularly older) children understood the meaning of these cues (e.g., “their muscles are not moving well”; age 8.8, “she has a big disability”; age 9.3), others seemed unaware (e.g., for wheelchair images: “watching a movie”; age 4.5, “relaxing in a chair”; age 6.6, “I think they’re in a plane”; age 5.3). Children also diverged in their intuitions regarding the source of perceived impairments (e.g., providing accident-accounts such as “his leg got cut off” (age 8.8), vs. congenital accounts such as “he was born without a leg”; age 6.25). Taken together, the open impression responses provided a fascinating glimpse into children’s developing understanding of the meaning of cues to disability

and foreignness, the range of explanations utilized to address perceived anomalies in speech and appearance and the ensuing generalizing inferences drawn.

4.3.Research questions

For each of the primary research questions considered below (conceptual differentiation, disability spread, prejudice), relevant descriptive statistics, results of mixed-effects models, *post hoc* tests and visualizations are provided, followed by narrative explication of the findings.

4.3.1. Conceptual differentiation:

4.3.1.1.To what extent will children selectively associate physical disability cues (i.e., amputation, wheelchair) but not foreignness cues (i.e., foreign garb) with reduced physical competence (running ability)?

Table 36. Descriptive Statistics for competence (running) scores on 3-point Likert scale.

Competence Scores (Visual)		
Category	Mean	SD
Typical	2.58	0.68
Amputee	1.29	0.63
Wheelchair	1.43	0.72
Foreign	2.11	0.78

Table 37. Coefficient estimates from the mixed effects model predicting physical competence scores from image category (Note: North American-presenting images without disability are the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	2.58	0.08	30.88	< .0001
Amputee	-1.29	0.12	-10.97	< .0001
Wheelchair	-1.15	0.12	-9.79	< .0001
Foreign	-0.47	0.12	-4.01	< .0001

Table 38. Post hoc pairwise comparisons assessing differences in physical competence scores between image categories

contrast	estimate	*SE*	t	p
Typical - Amputee	1.29	0.12	10.97	< .0001
Typical - WC	1.15	0.12	9.79	< .0001
Typical - Foreign	0.47	0.12	4.01	< .0001
Amputee - WC	-0.14	0.12	-1.18	0.99
Amputee - Foreign	-0.82	0.12	-6.96	< .0001
WC - Foreign	-0.68	0.12	-5.78	< .0001

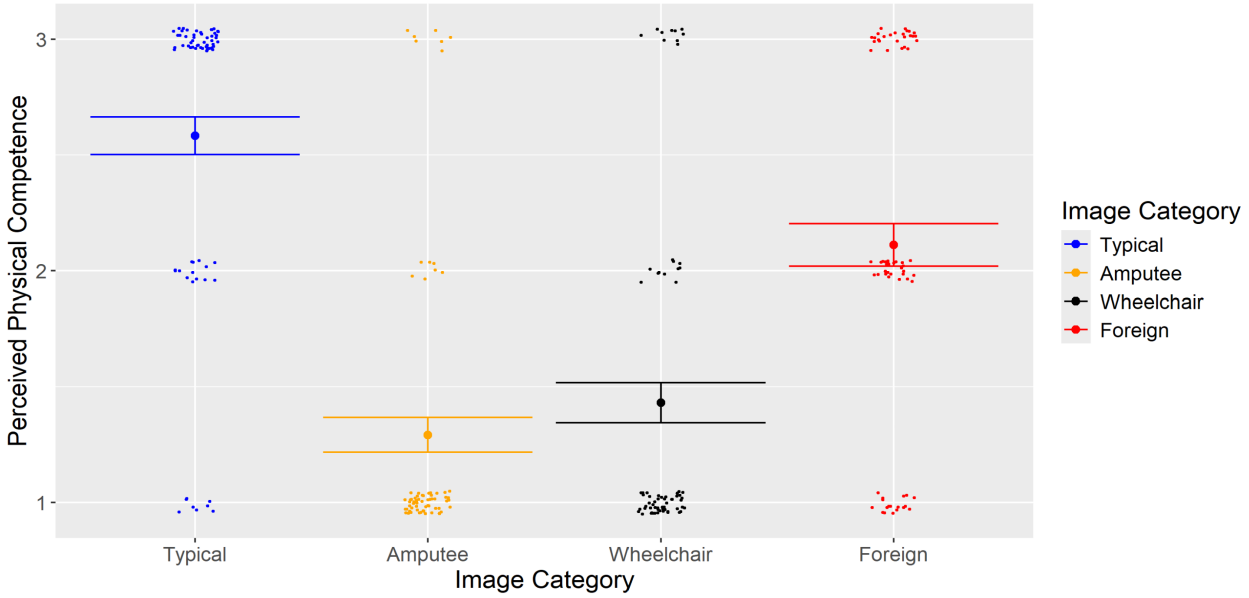

Figure 9. **Perceived Physical Competence scores across the four image categories. "Typical" refers to North American presenting images without disability, "Amputee" refers to North American presenting images with amputation, "Wheelchair" refers to North American presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Error bars represent standard errors around the mean.**

For the visual stimuli, participants clearly provided significantly higher physical competence scores for non-disabled North American-presenting and foreign-presenting images in comparison to amputee and wheelchair images (see Tables 36-38 and Figure 9). In addition, foreign-presenting images received significantly lower physical competence scores than typical images but significantly higher physical competence scores than the two image categories depicting disabilities (see Table 38 and Figure 9). Thus, children clearly selectively attributed low physical competence to the disability images, indicating that they differentiated between foreignness and disability cues *vis-à-vis* physical competence perceptions, as predicted.

4.3.1.2.To what extent will children selectively associate foreignness cues (i.e., foreign accent and garb) but not disability cues (i.e., amputation, wheelchair) with being from another country?

Table 39. Descriptive statistics and binomial tests for foreignness judgments (binary choice: "America" or "Another country far away")

Foreignness Judgments (Visual)				Foreignness Judgments (Speech)		
Category	% "far away" selection	Two-tailed Binomial significance test		Category	% "far away" selection	Two-tailed Binomial significance test
Typical	37.5%	$p = 0.04$		L1	43.6%	$p = 0.34$
Amputee	43%	$p = 0.28$		L2	69%	$p = 0.001$
Wheelchair	50%	$p = 0.99$		ASD	55%	$p = 0.47$
Foreign	73%	$p = .00007$				

Table 40. Coefficient estimates from the mixed effects model predicting foreignness judgments from image category, age, and their interaction (Note: North American-presenting images without disability are the reference group).

	Estimate	*SE*	*z*	*p*
(Intercept)	-0.55	0.26	-2.12	0.034
Amputee	0.04	0.4	0.11	0.9
Wheelchair	0.58	0.36	1.61	0.1
Foreign	1.56	0.4	4.002	0.0006
Age	-0.40	0.28	-1.41	0.15
Amputee x age	-0.85	0.5	-1.74	0.08
Wheelchair x age	0.02	0.38	0.07	0.94
Foreign x age	0.85	0.42	2.02	0.04

Table 41. Post hoc trend analysis assessing age related change in foreignness judgments across image categories

Image type	Age trend	*SE*	asymp.LCL	asymp.UCL	*z*	*p*
Typical	-0.40	0.28	-0.95	0.15	-1.42	0.16
Amputee	-1.26	0.40	-2.04	-0.47	-3.14	<0.001
WC	-0.37	0.26	-0.88	0.14	-1.42	0.15
Foreign	0.46	0.32	-0.16	1.08	1.44	0.15

Table 42. Post hoc pairwise comparisons assessing differences in foreignness judgments across image categories

contrast	estimate	*SE*	*z*	*p*
Typical - Amputee	-0.05	0.40	-0.12	0.99
Typical - WC	-0.59	0.36	-1.61	0.372
Typical - Foreign	-1.57	0.39	-4.00	0.0004
Amputee - WC	-0.54	0.39	-1.38	0.99

contrast	estimate	*SE*	*z*	*p*
Amputee - Foreign	-1.52	0.42	-3.65	0.0015
WC - Foreign	-0.98	0.38	-2.55	0.053

Table 43. Coefficient estimates from the logistic mixed effects model predicting foreignness judgments from speech category, age, and their interaction (Note: Neurotypical North American English speech is the reference group)

	Estimate	*SE*	*z*	*p*
(Intercept)	-0.26	0.28	-0.95	0.34
L2	1.05	0.38	2.717	0.006
ASD	0.39	0.37	1.04	0.29
Age	-0.98	0.34	-2.85	0.004
L2 x age	1.07	0.44	2.43	0.015
ASD x age	0.8	0.42	1.87	0.061

Table 44. Post hoc trend analysis assessing age related change in foreignness judgments across speech categories

Speech Category	Age trend	*SE*	asymp.LCL	asymp.UCL	*z*	*p*
L1	-0.98	0.34	-1.66	-0.31	-2.86	<0.001
L2	0.09	0.27	-0.45	0.62	0.31	0.75
ASD	-0.19	0.25	-0.68	0.31	-0.73	0.46

Table 45. Post hoc pairwise comparison assessing differences in foreignness judgments across speech categories

Contrast	Estimate	*SE*	*z*	*p*
L1 - L2	-1.06	0.39	-2.72	0.02
L1 - ASD	-0.39	0.38	-1.04	0.89
L2 - ASD	0.66	0.37	1.80	0.22

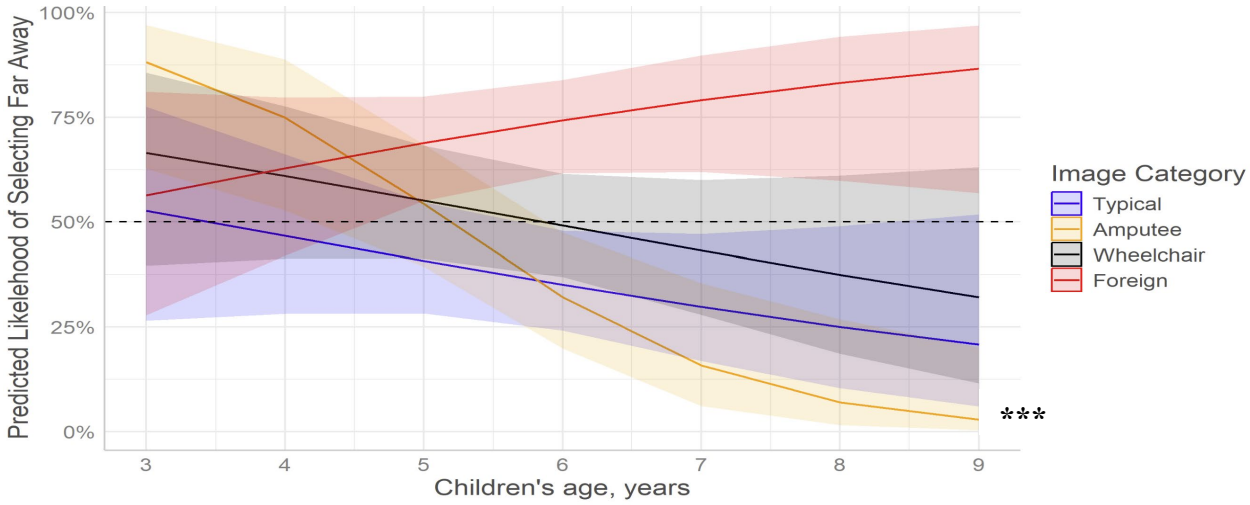

Figure 10. **Relationship between participants' age (in years) and predicted likelihood of selecting "far away". "Typical" refers to North American-presenting images without**

disability, "Amputee" refers to North American-presenting images with amputation, "WC" refers to North American-presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Lines represent the predicted likelihood of selecting "far away" for each image category in relation to participants' age. Shaded areas around the lines represent confidence intervals based on the standard errors from the logistic regression. Asterisks represent significant age trends (* for $p < 0.001$, ** for $p < 0.01$, * for $p < 0$ .05).**

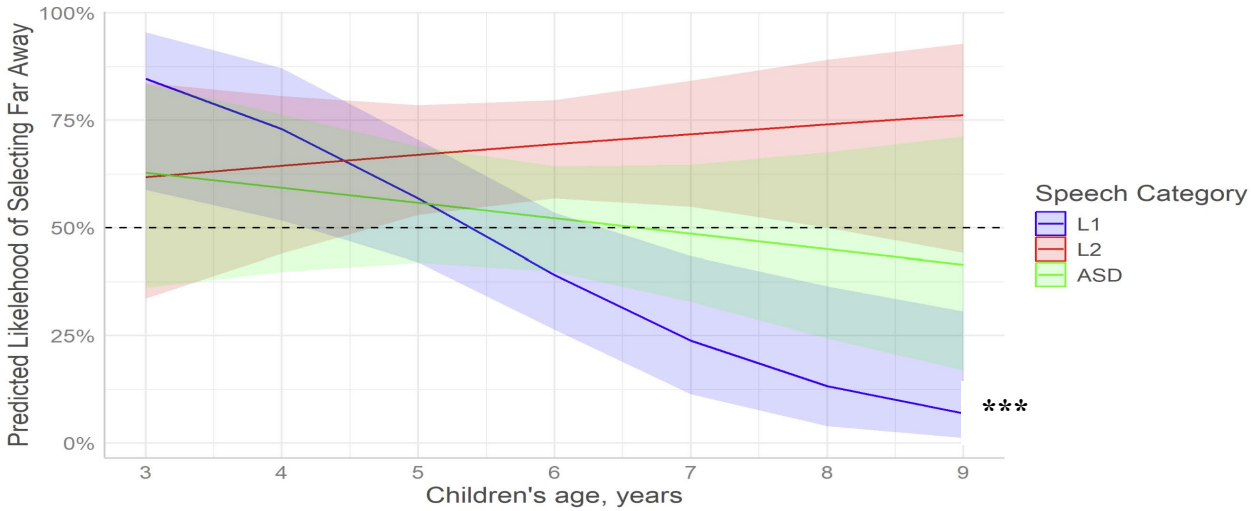

Figure 11. **Relationship between participants' age (in years) and predicted likelihood of selecting "far away". L1 refers to neurotypical North American English speakers, L2 refers to neurotypical Spanish speakers learning English as a second language, ASD refers to North American English speakers with autism spectrum disorder. Lines represent the predicted likelihood of selecting "far away" for each speech category in relation to participants' age. Shaded areas around the lines represent confidence intervals based on the standard errors from the logistic regression. Asterisks represent significant age trends. (*** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0$ .05).**

For the visual stimuli, children as a group clearly reserved "far away" ratings for foreign-presenting images in comparison to all three of the other image categories, as predicted (see Table 39 and Figure 10). However, unexpectedly, the difference between foreign and wheelchair images was only marginally significant (Table 42), reflecting some level of uncertainty regarding the perceived foreignness of the wheelchair images. Along those lines, whereas there was a significant age-related increase in children's tendency to associate amputee images with being

from America, the age-related change in children's foreignness judgments for the wheelchair images was not significant (though it was trending towards ratings of "America" – Table 41). Finally, exploratory analysis revealed that male (but not female) foreign garb images were the ones that reliably elicited foreignness judgements to a greater extent than the other three image categories (see supplementary materials, section 4).

For speech stimuli, children as a group clearly identified L2 speakers as more likely to be from "far away" than L1 speakers but seemed relatively unclear as to ASD speakers' geographical origin in this regard (see Figure 11). With age, children showed an increase in identifying L1 speakers as "from America" but showed no significant age-related changes in foreignness judgments provided to L2 and ASD speakers (see Table 44 and Figure 11). Taken together, children reliably selectively associated foreign garb and accent with being from another country far away. They also reliably selectively associated typical images without disability (regardless of age), as well as amputee images and L1 speakers (particularly for older children) with being from America. On the other hand, children seemed unsure as to whether WC images and ASD speakers were "From America" or "another country far away".

To shed light on the children who were unsure of the foreignness status of WC images and ASD speakers, we explored moderating factors. For the visual stimuli, this moderation analysis revealed that children whose parents had higher levels of education as well as children with exposure to other children with a disability were more likely to indicate that WC images were "from America" (see Tables 46 and 47 below). In contrast, for speech, none of the moderators were reliably associated with an increased tendency to associate ASD speakers with being "from America"

Table 46. Post hoc trend analysis assessing the effect of parental education on the likelihood of selecting typical images for each image category

Image type	Parental education trend	*SE*	asymp.LCL	asymp.UCL	*z*	*p*
Typical	0.1603568	0.24920	-0.32808	0.6487	0.6434	0.51992
Amputee	0.269054	0.253072	-0.22695	0.76506	1.0631	0.28771
WC	-0.498129	0.263757	-1.015084	0.01882	-1.8885	0.05894
Foreign	-0.190532	0.277133	-0.733704	0.35263	-0.6875	0.49175

Table 47. Post hoc pairwise comparisons assessing the effect of disability exposure (child) on the likelihood of selecting typical images for each image category

contrast	Disability exposure (child)	estimate	*SE*	*z*	*p*
Typical - Amputee	no	-0.25131	0.40986	-0.6131	0.927933
Typical - WC	no	-0.502628	0.41147	-1.22152	0.61311
Typical - Foreign	no	-0.944461	0.42238	-2.23604	0.11358
Amputee - WC	no	-0.25131	0.409865	-0.61316	0.927933
Amputee - Foreign	no	-0.693147	0.420812	-1.64716	0.352112
WC - Foreign	no	-0.441832	0.422381	-1.04605	0.722279
Typical - Amputee	yes	-1.2133	0.899737	-1.34851	0.99
Typical - WC	yes	-0.93430	0.813841	-1.14802	0.65967
Typical - Foreign	yes	-3.55534	0.985609	-3.60725	0.001759
Amputee - WC	yes	-0.9343	0.813842	-1.14802	0.659676

Amputee - Foreign	yes	-3.55534	0.985610	-3.60725	0.001759
WC - Foreign	yes	-2.62103	0.90787	-2.88699	0.020314

4.3.1.3.Conceptual Differentiation Summary

Children selectively associated illustrations featuring visible disability cues (i.e., amputation, wheelchair) with reduced physical competence in comparison to illustrations featuring non-disabled North American and foreign-presenting children. Furthermore, age was not a meaningful predictor in the model predicting physical competence scores, indicating that this differentiation was stable across ages (that is, for those who passed the reliability check). This suggests that cues to physical competence, such as our images of wheelchairs or crutches are relatively salient to young children, consistent with prior research (e.g., Diamond and Hestenes, 1996). In addition, children reliably selectively associated foreign garb and foreign accent with being from another country far away, while typical images without disability (regardless of age), as well as amputee images and L1 speakers (particularly for older children) were associated with being "from America". Thus, there was evidence suggesting that children in this study selectively differentiated between multimodal cues to foreignness and disability, as predicted.

On the other hand, children seemed unsure as to whether wheelchair images and ASD speakers were "from America" or "another country far away" and there were no significant age-related changes in this regard (though there was a trending age-related increase in the tendency to associate wheelchair images with being "from America"). This suggests that children may have struggled to provide a satisfactory explanatory account for the perceptual variability detected in these stimuli. This pattern also highlights potential differences in children's reactions

to amputee images versus. the wheelchair images that contained multiple disability cues (i.e., contracture in the wrist and torticollis in the neck, see McCoy and Banks, 2012).

The fact that foreign illustrations received lower physical competence scores than (non-disabled) North American illustrations can be interpreted in several ways. On the one hand, it is possible that children's foreignness concepts entail perceptions of reduced physical competence (though to a lesser extent than their disability concepts). On the other hand, the specific cues used to convey foreignness in our illustrations (traditional Spanish flamenco dress for girls and a Matador's traje de luces for boys) may have inadvertently communicated reduced athletic prowess and/or restrictive clothing in comparison to the more athletic outfits (e.g., sneakers and shorts) featured in the other illustrations. Furthermore, the mean physical competence scores provided for foreign images (2.11) were above the median score (2), which highlights that the foreign/typical difference effect was due to typical images receiving particularly high physical competence scores rather than foreign images receiving low scores. While future research is needed, the fact that children did not provide lower physical competence scores for L2 speakers suggests that the visual stimuli per-se – rather than children's foreignness concepts – may have driven the observed lower physical competence scores for foreign images.

Taken together, the emerging conclusion regarding children's conceptual differentiation was mixed. On the one hand, children reliably associated low physical competence with images depicting physical disabilities and reliably associated foreignness with foreign garb and accent. On the other hand, children did not reliably associate wheelchair images and ASD speakers with being from America. Children's difficulty associating these stimuli categories with being "from America" points to the protracted nature of children's developing tolerance for within-group diversity as well as their gradually developing understanding of the diverse manifestations of

disabilities. Finally, the moderating effect of parental education and disability exposure points to malleability in this regard.

4.3.2. Disability spread:

4.3.2.1. To what extent will children selectively associate disability cues (i.e., disordered speech, amputation, wheelchair) but not foreignness cues (i.e., foreign accent and garb) with greater dependence?

Table 48. Descriptive statistics for dependence (needing-help) scores on 3-point Likert Scale

Dependence Scores (Visual)				Dependence Scores (Speech)		
Category	Mean	SD		Category	Mean	SD
Typical	1.65	0.84		L1	1.90	0.83
Amputee	2.52	0.75		L2	2.11	0.85
Wheelchair	2.54	0.75		ASD	2.09	0.84
Foreign	1.61	0.84				

Table 49. Coefficient estimates from the mixed effects model predicting perceived dependence scores from image category, age, and their interaction (Note: North American-presenting images without disability are the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	1.69	0.10	16.95	< .0001
Amputee	0.80	0.13	6.10	< .0001
Wheelchair	0.83	0.13	6.33	< .0001
Foreign	-0.08	0.13	-0.59	.558
Age	-0.23	0.10	-2.29	.023
Amputee x Age	0.44	0.13	3.38	.001
Wheelchair x Age	0.26	0.13	1.95	.052
Foreign x Age	0.06	0.13	0.44	.664

Table 50. Post hoc trend analysis assessing age-related change in dependence scores between image categories

Image type	Age trend	*SE*	lower.CL	upper.CL	*t*	*p*
Typical	-0.23	0.10	-0.43	-0.03	-2.29	0.02
Amputee	0.22	0.10	0.02	0.41	2.15	0.03
WC	0.03	0.10	-0.17	0.22	0.28	0.78
Foreign	-0.17	0.10	-0.37	0.03	-1.72	0.09

Table 51. Coefficient estimates from the mixed effects model predicting perceived dependence scores from speech category, age, and their interaction (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	1.91	0.11	18.06	< .001
L2	0.19	0.12	1.59	.114
ASD	0.23	0.12	1.99	.049
Age	-0.25	0.11	-2.34	.020
L2 x Age	0.09	0.12	0.73	.465
ASD x Age	0.16	0.12	1.37	.173

Table 52. Post hoc trend analysis assessing age-related change in dependence scores between speech categories

Speech Category	Age trend	*SE*	lower.CL	upper.CL	*t*	*p*
L1	-0.25	0.11	-0.46	-0.04	-2.34	0.02
L2	-0.16	0.11	-0.37	0.05	-1.53	0.13
ASD	-0.09	0.11	-0.30	0.12	-0.81	0.42

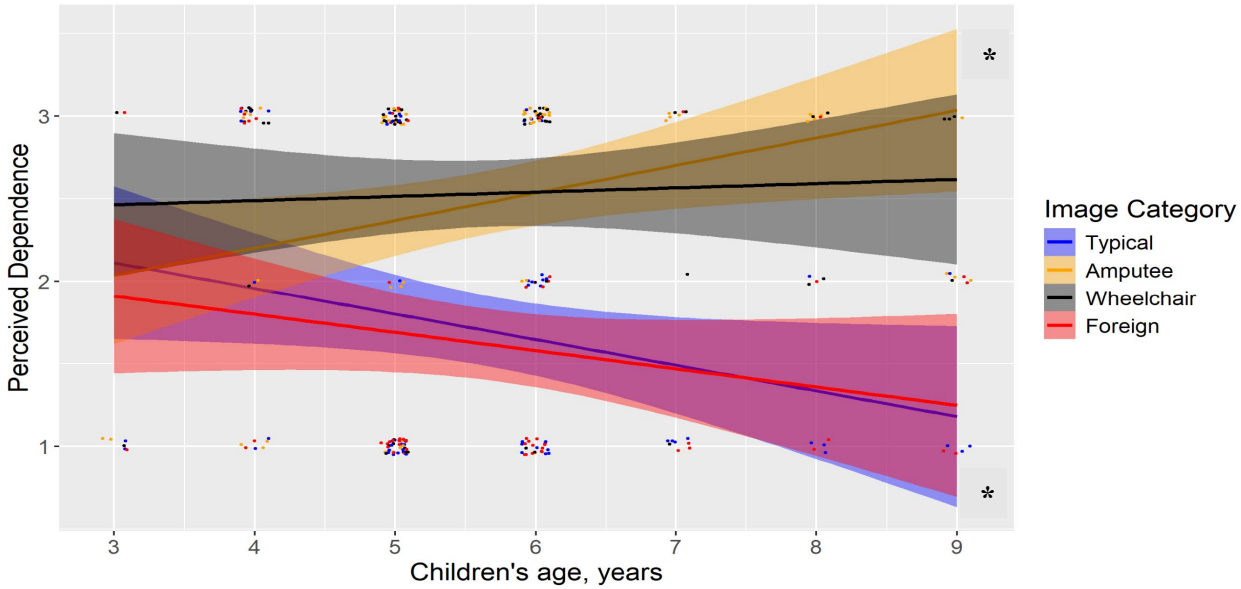

Figure 12. **Perceived dependence scores across the four image categories in relation to participants' age. "Typical" refers to North American presenting images without disability, "Amputee" refers to North American presenting images with amputation, "Wheelchair" refers to North American presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Lines represent age related trends in perceived dependence scores for each image category. Shaded areas around the lines represent confidence intervals based on the standard errors from the mixed-effects regression. Asterisks represent significant age trends (*** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0$ .05).**

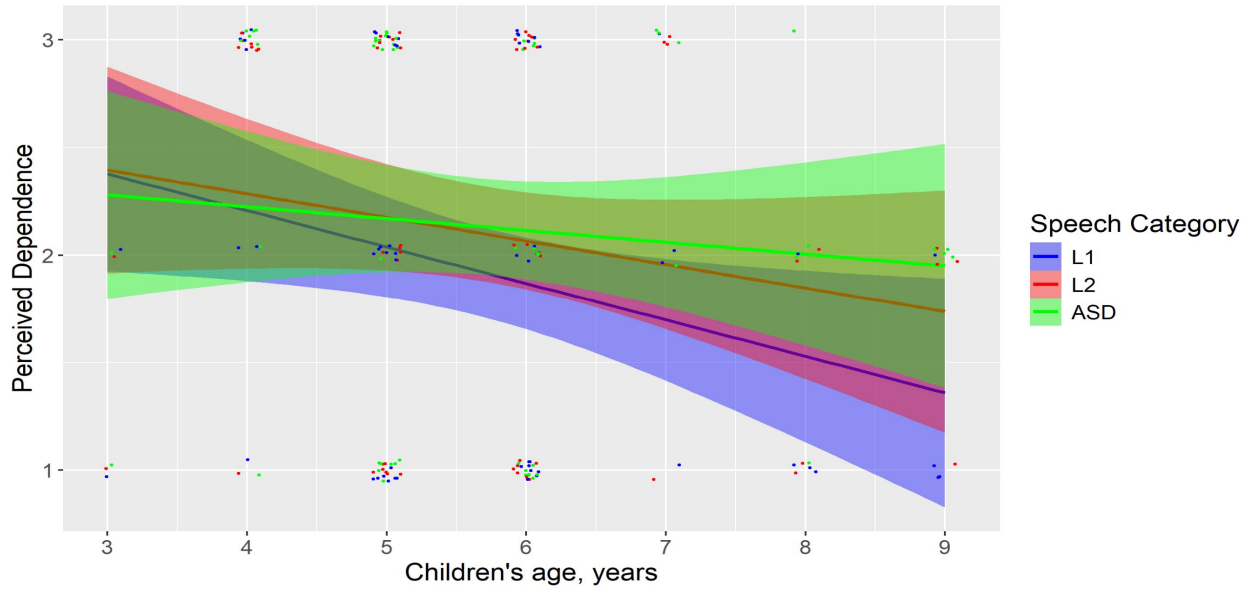

Figure 13. **Perceived dependence scores across the three speech categories in relation to participants' age. L1 refers to neurotypical North American English speakers, L2 refers to neurotypical Spanish speakers learning English as a second language, ASD refers to North American English speakers with autism spectrum disorder. Lines represent age related trends in perceived dependence scores for each image category. Shaded areas around the lines represent confidence intervals based on the standard errors from the mixed-effects regression. Asterisks represent significant age trends (*** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0$ .05).**

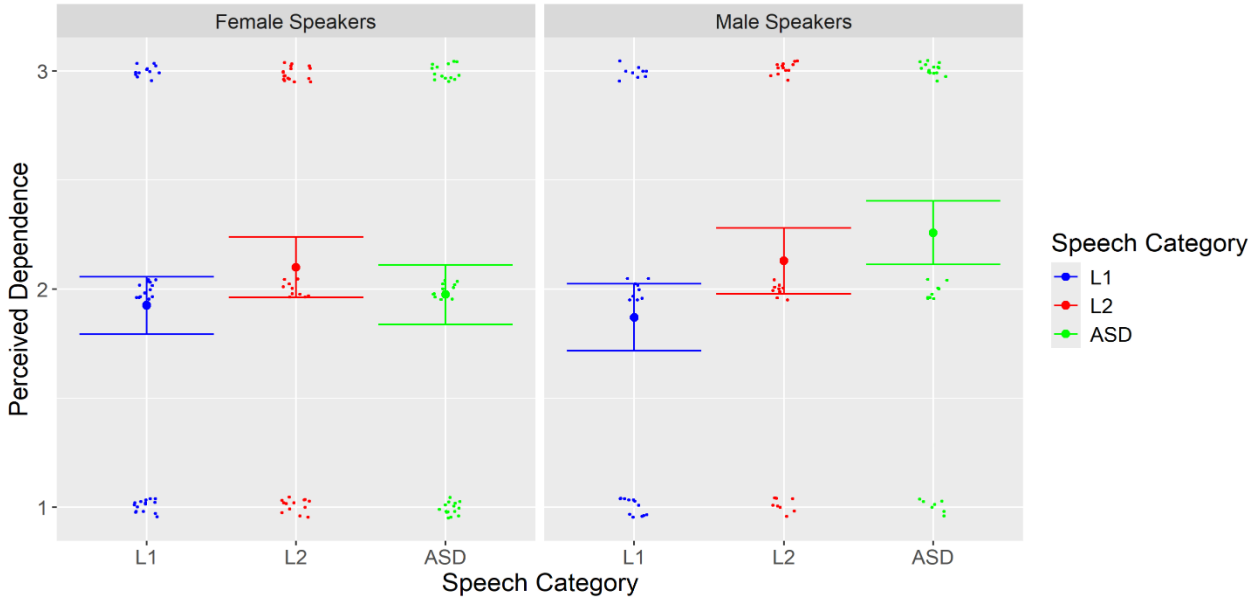

Figure 14. **Perceived dependence scores across the three speech categories in relation to speaker's gender. L1 refers to neurotypical North American English speakers, L2 refers**

to neurotypical Spanish speakers learning English as a second language, ASD refers to North American English speakers with autism spectrum disorder. Error bars represent standard errors around the mean.

For the visual stimuli, participants clearly provided significantly lower dependence scores for typical North American-presenting and foreign-presenting images in comparison to amputee and wheelchair images (see Tables 48-49 and Figure 12). Whereas older children became increasingly likely to associate typical images with lower dependence (and trended that way for foreignness judgments, as well), older children also became increasingly likely to associate amputee images with greater dependence (Table 50 and Figure 12). In contrast, there was no systematic age-related change in children's dependence ratings *vis-à-vis* the wheelchair images (see Table 50). All in all, there was clear evidence of disability spread in young children's evaluations of individuals with visual disability cues.

For the speech stimuli, there was a trending tendency to associate greater dependence scores with ASD speakers in comparison to L1 speakers (Table 51), but age-related change was only observed for evaluation of L1 speakers (Table 52). Namely, with age, children became increasingly likely to associate lower dependence with L1 speakers. Thus, there was some evidence of disability spread in young children's evaluations of ASD speakers, but the magnitude of this effect was smaller in comparison to the pattern observed for the visual cues to physical disability. Lastly, exploratory analysis revealed that the tendency to associate greater dependence with ASD speakers was only evident for male (but not female) speakers with ASD (see Figure 14 and supplementary materials, section 4).

4.3.2.2.To what extent will children selectively associate reduced physical competence (running ability) with ASD speakers but not L2 speakers?

Table 53. Descriptive Statistics for physical competence (running) scores on 3-point Likert scale.

Competence Scores (Speech)		
Category	Mean	SD
L1	2.32	0.84
L2	2.15	0.82
ASD	2.29	0.79

Table 54. Coefficient estimates from the mixed effects model predicting physical competence scores from speech category (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	2.32	0.10	23.84	< .001
L2	-0.17	0.12	-1.40	.163
ASD	-0.03	0.12	-0.23	.816

Table 55. Post hoc pairwise comparisons assessing differences in physical competence scores between speech categories

contrast	estimate	*SE*	*t*	*p*
L1 - L2	0.17	0.12	1.40	0.49
L1 - ASD	0.03	0.12	0.23	0.99
L2 - ASD	-0.14	0.12	-1.17	0.73

In contrast to the dependence scores for the speech stimuli which revealed a tendency to associate ASD speakers, but not L2 speakers, with greater dependence, participants did not display significant differences in their judgments of perceived physical competence across the three speech conditions (see Tables 53-55). Thus, we found no evidence to suggest that children, at any age, attributed reduced physical competence to speakers with ASD, contrary to our predictions.

4.3.2.3.Disability Spread Summary

As noted, we operationalized spreading in two ways. First, we assessed the extent to which participants would selectively rate speakers with ASD with reduced physical competence

and found no evidence for such a pattern. Second, we assessed the degree to which participants would selectively rate speakers with ASD as well as illustrations featuring physical disabilities with increased dependence. For speech, there was a trending tendency for children to associate enhanced dependence with ASD speakers. More robustly, children clearly associated enhanced dependence with images depicting physical disabilities. Notably, while this tendency became increasingly apparent for older children in the case of amputees, there was no age-related change in the case of wheelchairs. Thus, the perception that the wheelchair images were highly dependent was more apparent for young children in comparison to amputee images. In sum, there was some evidence of disability spread in the case of ASD speakers and robust evidence for disability spread in the case of physical disabilities.

It is worth considering the findings regarding disability spread in conjunction with the conceptual differentiation findings reported above. Specifically, the fact that children were relatively unsure regarding the foreignness of ASD speakers suggests that children who participated in this study were still in the midst of constructing perceptual/conceptual differentiation between foreignness cues and disability cues. The difficulty that children showed in making these discriminations would be expected to have implications for other phenomena of interest, such as disability spread. Concretely, since children were relatively unsure whether to attribute the atypical features they detected in ASD speech to foreignness or disability, it is somewhat predictable that we would observe limited evidence of disability spread in the case of ASD speakers. Nevertheless, the trending tendency to associate (male) ASD speakers with enhanced dependence points to the early emergence of disability spread tendencies in children's evaluations of ASD speakers.

4.3.3. Peer Rejection

4.3.3.1. To what extent will children express reduced interest in pursuing friendships with foreign or disabled targets (as indicated by speech and/or visual appearance) in comparison to non-foreign, non-disabled targets?

Table 56. Descriptive Statistics for friendship scores on 3-point Likert Scale

Friendship Scores (Visual)				Friendship Scores (Speech)		
Category	Mean	SD		Category	Mean	SD
Typical	2.04	0.88		L1	2.01	0.88
Amputee	1.86	0.86		L2	1.88	0.87
Wheelchair	1.75	0.80		ASD	1.87	0.82
Foreign	1.72	0.80				

Table 57. Coefficient estimates from the mixed effects model predicting friendship scores from image category (Note: North American-presenting images without disability are the reference group).

	Estimate	*SE*	*t*	*p*
Intercept	2.04	0.10	20.67	< .001
Amputee	-0.18	0.11	-1.68	.095
Wheelchair	-0.29	0.11	-2.71	.007
Foreign	-0.32	0.11	-2.97	.003

Table 58. Pair-wise post hoc analysis assessing differences in friendship scores between the image categories

contrast	Estimate	*SE*	*t*	*p*
Typical - Amputee	0.18	0.11	1.68	0.57
Typical - WC	0.29	0.11	2.71	0.04
Typical - Foreign	0.32	0.11	2.97	0.02
Amputee - WC	0.11	0.11	1.03	0.99
Amputee - Foreign	0.14	0.11	1.29	0.99
WC - Foreign	0.03	0.11	0.26	0.99

Table 59. Coefficient estimates from the mixed effects model predicting friendship scores from speech category (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	2.014	0.10	19.694	< 0.0001
L2	-0.127	0.12	-1.04	0.3
ASD	-0.14	0.12	-1.15	0.25

Table 60. Pair-wise post hoc analysis assessing differences in friendship scores between the speech categories

contrast	Estimate	*SE*	*t*	*p*
L1 - L2	0.127	0.12	1.04	0.55
L1 - ASD	0.1408	0.12	1.156	0.48
L2 - ASD	0.0141	0.12	0.116	0.99

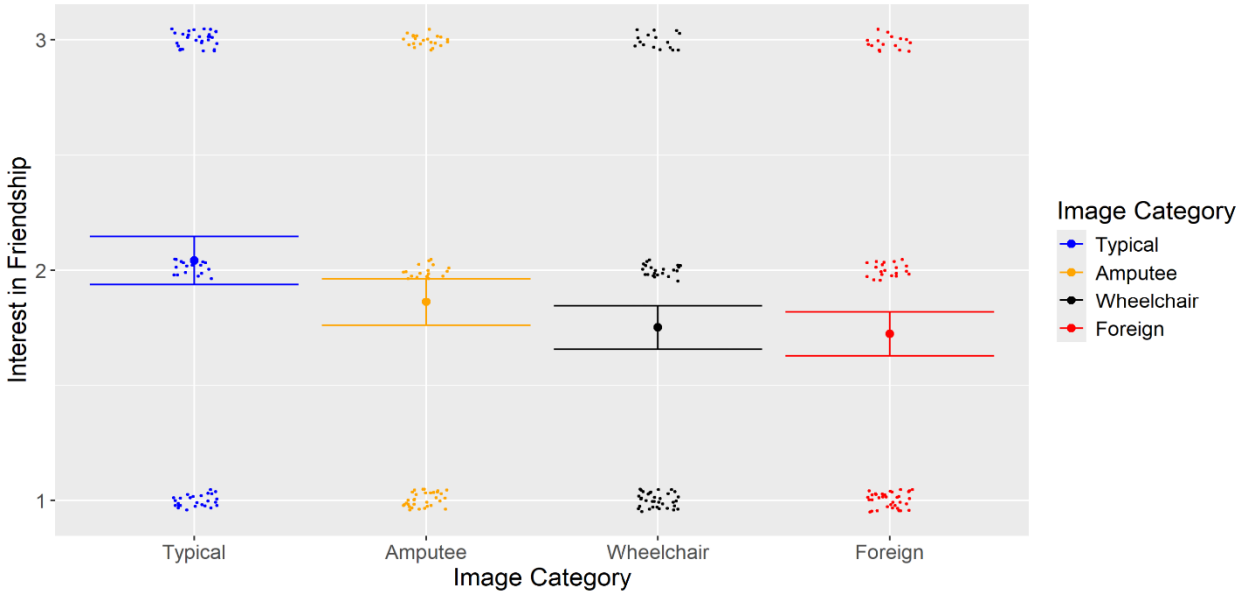

Figure 15. **Friendship scores across the four image categories. "Typical" refers to North American presenting images without disability, "Amputee" refers to North American presenting images with amputation, "Wheelchair" refers to North American presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Error bars represent standard errors around the mean.**

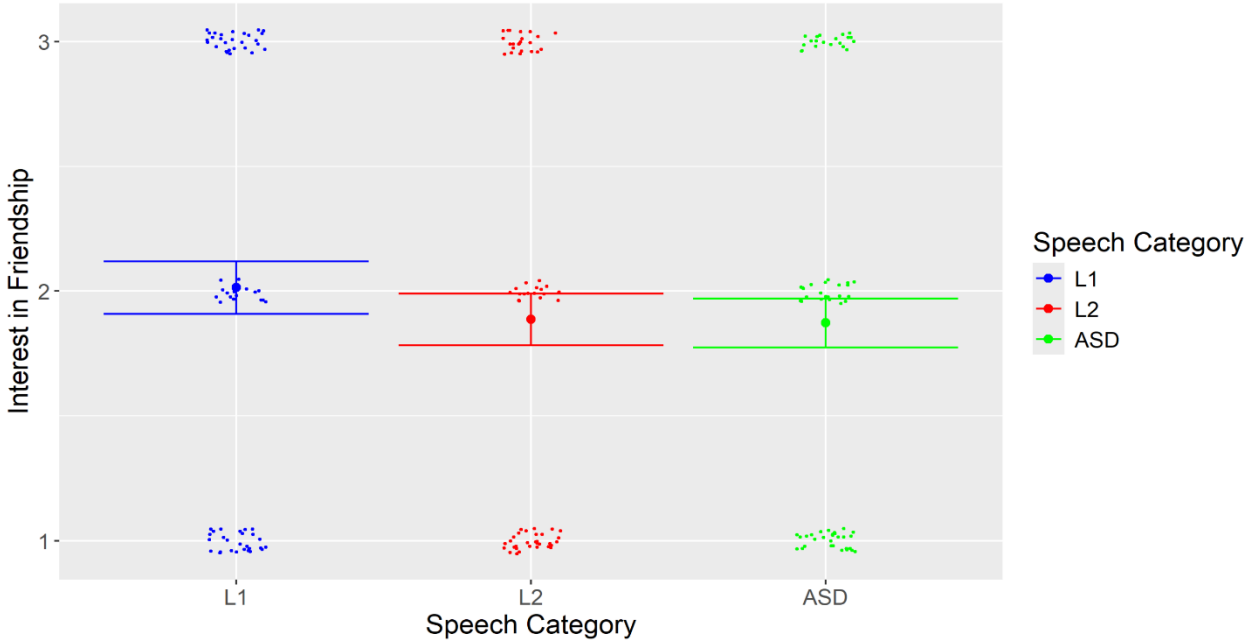

Figure 16. **Friendship scores across the three speech categories. L1 refers to neurotypical North American English speakers, L2 refers to neurotypical Spanish speakers learning English as a second language, ASD refers to North American English speakers with autism spectrum disorder. Error bars represent standard errors around the mean.**

For the visual stimuli, participants were significantly less interested in pursuing friendships with children depicted in wheelchair and foreign (but not amputee) images in comparison to typical images (see Table 56-58 and Figure 15). In contrast to the visual stimuli, we did not observe significant differences in friendship scores between speech conditions (Tables 59-60 and Figure 16). In sum, children were more likely to express peer rejection when evaluating images in comparison to speakers. When peer rejection tendencies were detected, they were primarily against foreign and WC images (but not amputee images).

The null results for the friendship ratings of speakers raise questions regarding previous characterizations of speech as a particularly potent driver for social cognition (See Kinzler, 2021 for discussion), at least in the case of children. However, since the audio versus visual stimuli were not systematically balanced and scaled in terms of dimensions such as foreignness or

disability salience, strong conclusions regarding the implications of the observed differential response patterns to these stimuli vis-à-vis friendship scores are not warranted and future investigation is needed (see general discussion for further elaboration).

As noted, whereas significant peer rejection was not observed in the case of speech, participants did express reduced interest in pursuing friendships with children depicted in wheelchairs with multiple cues to a disability (e.g., contracture in the wrist and torticollis in the neck) and foreign images in comparison to non-foreign, non-disabled illustrations. Notably, this was not the case for images depicting children with an amputated limb, thus conceptually replicating previous research with adult participants (Gordon, Minnes and Holden, 1990). This raises the possibility that young children may attribute different meaning to a disability perceived as localized or peripheral (e.g., child on crutches with amputation) in comparison to a disability perceived as central (e.g., wheelchair bound child with signs of cerebral palsy). In turn, beyond greater prejudice, such attributions may also be associated with other factors such as greater inductive potential (a topic explored in study 2). On the other hand, since friendship scores for wheelchair images were roughly identical to foreign images, blatant disability prejudice that extended beyond an inclination to avoid friendship with those in foreign garb was not detected.

4.3.3.2. *Is peer rejection most closely associated with ratings of physical competence, dependence, or comprehensibility (the latter in the case of speech)?*

4.3.3.2.1. Predictors of Prejudice – Visual Stimuli

To assess factors contributing to the expression of peer rejection, we explored relationships between perceived physical competence, dependence, and friendship scores for the visual stimuli. Notably, exploratory analysis revealed that participants' gender significantly moderated the relationship between friendship scores and perceived dependence. In contrast,

gender was not a significant moderator of the relationship between perceived competence and friendship scores and was thus left out of the competence model. The results of these mixed effects models and *post hoc* trend analyses are provided below.

Table 61. Coefficient estimates from the mixed effects model predicting friendship scores from the image type, physical competence scores, and their interaction (Note: North American-presenting images without disability are the reference group).

	Estimate	SE	t	p
(Intercept)	1.09	0.33	3.27	.001
Amputee	0.84	0.38	2.23	.026
WC	0.67	0.39	1.74	.083
Foreign	0.54	0.39	1.37	.171
Competence	0.37	0.12	3.00	.003
Amputee x Competence	-0.42	0.18	-2.33	.020
WC x Competence	-0.37	0.18	-2.13	.034
Foreign x Competence	-0.33	0.16	-2.05	.041

Table 62. Post hoc trend analysis assessing the effect of perceived physical competence on friendship scores for each image category

Image_type	Competence Trend	*SE*	lower.CL	upper.CL	*t*	*p*
Typical	0.37	0.12	0.13	0.61	2.99	<0.0001
Amputee	-0.05	0.13	-0.32	0.21	-0.41	0.68
WC	-0.01	0.12	-0.23	0.22	-0.04	0.97
Foreign	0.04	0.11	-0.17	0.26	0.41	0.68

Table 63. Coefficient estimates from the mixed effects model predicting friendship scores from the image type, dependence scores, participants gender and their interaction (Note: Female participants assessing North American-presenting images without disability are the reference group).

	Estimate	SE	t	p
(Intercept)	2.36	0.27	8.70	< .001
Amputee	-0.50	0.58	-0.86	.389
WC	-1.70	0.49	-3.44	.001
Foreign	-0.33	0.33	-0.99	.322
Dependence	-0.31	0.15	-2.06	.041
Male	-0.21	0.39	-0.54	.590
Amputee x Dependence	0.31	0.24	1.27	.204
WC x Dependence	0.71	0.23	3.14	.002
Foreign x Dependence	0.07	0.19	0.34	.731
Amputee x Male	0.59	0.73	0.81	.419
WC x Male	1.47	0.70	2.10	.037
Foreign x Male	0.21	0.49	0.43	.665

	Estimate	SE	t	p
Dependence x Male	0.33	0.20	1.62	.107
Amputee x Dependence x Male	-0.49	0.31	-1.56	.119
WC x Dependence x Male	-0.76	0.31	-2.49	.013
Foreign x Dependence x Male	-0.24	0.27	-0.88	.381

Table 64. Post hoc trend analysis assessing the effect of perceived dependence on friendship scores for each image category in relation to participants' gender

Image Category	Gender	Dependence Trend	*SE*	lower.CL	upper.CL	*t*	*p*
Typical	Female	-0.31	0.15	-0.61	-0.01	-2.05	0.04
Amputee	Female	0.00	0.19	-0.38	0.38	0.00	1.00
WC	Female	0.40	0.16	0.09	0.71	2.55	0.01
Foreign	Female	-0.24	0.14	-0.53	0.04	-1.69	0.09
Typical	Male	0.02	0.14	-0.25	0.29	0.15	0.88
Amputee	Male	-0.16	0.14	-0.44	0.12	-1.11	0.27
WC	Male	-0.03	0.16	-0.35	0.29	-0.20	0.84
Foreign	Male	-0.15	0.14	-0.42	0.12	-1.08	0.28

For the visual images, the association between perceived physical competence and friendship scores varied significantly across the image categories (Table 61). Namely, perceived physical competence scores were significantly associated with greater friendship scores but only for the typical images (Table 62). Thus, perceptions of reduced physical competence detected in images depicting physical disabilities (or foreign images) were not related to children's friendship scores provided to these images. In fact, when controlling for physical competence scores (which were selectively associated with friendship ratings provided to typical images), participants actually favored the amputee images (and, to a lesser extent, the WC images as well) over the typical images (see Table 61).

In contrast to perceptions of physical competence which were not associated with friendship scores provided to images depicting disabilities, perceptions of greater dependence were significantly associated with greater friendship scores for female (but not male) participants' ratings of the wheelchair images, but not for the other images (Table 64). In fact, when controlling for dependence scores (which were selectively associated with greater

friendship scores provided by female participants to wheelchair images), there was a significant preference for typical images over wheelchair images, but not foreign or amputee images (see Table 63). This reflects the degree to which perceived dependence was uniquely related to friendship scores provided to wheelchair images.

4.3.3.2.2. Predictors of Peer Rejection – Speech Stimuli

To assess factors contributing to the expression of peer rejection, we explored the relationship between perceived physical competence, dependence, comprehensibility, and friendship scores for the speech stimuli.

First, we assessed the relationship between friendship and physical competence scores across the three speech conditions. Since a likelihood ratio test revealed that accounting for the interactions between speech category and competence scores did not significantly improve the model fit, we dropped the interaction term from the model. However, since we had an a priori interest in relative differences between speech categories vis-à-vis the effect of perceived physical competence on friendship scores, we also ran *post hoc* trend analyses exploring such differences. Results from a mixed-effects model predicting friendship scores from the main effects of speech category and competence scores are shown in Table 65. Results of the *post hoc* trend analyses are provided in Table 66.

Table 65. Coefficient estimates from the mixed effects model predicting friendship scores from audio type and perceived competence (Note: Neurotypical North American English speech – L1 – is the reference group).

	Estimate	SE	t	p
(Intercept)	1.56	0.19	8.13	< .001
L2	-0.09	0.12	-0.76	.450
ASD	-0.14	0.12	-1.10	.273
Competence	0.20	0.07	2.79	.006

Table 66. Post hoc trend analysis assessing the effect of perceived physical competence on friendship scores for each speech category

Speech Category	Competence Trend	*SE*	lower.CL	upper.CL	*t*	*p*
L1	0.28	0.11	0.05	0.50	2.42	0.02
L2	0.18	0.12	-0.05	0.41	1.54	0.12
ASD	0.11	0.12	-0.13	0.35	0.93	0.35

This model clarified that perceived physical competence, rather than speech category, significantly predicted (higher) friendship scores. Furthermore, *post hoc* analysis revealed that physical competence contributed to friendship scores primarily for L1 speakers (Table 66). Namely, perceived physical competence was not associated with friendship scores provided for ASD (or L2) speakers. However, since an interaction between competence and speech category was not warranted, these differences should be interpreted with caution.

Next, we repeated this analysis for perceived dependence. As was the case with perceived physical competence, model comparisons indicated that a model predicting friendship scores from the main effects of speech category and dependence, without an interaction term, provided the better fit. However, in contrast to perceived physical competence, model comparisons also revealed that adding dependence to a model predicting friendship scores from speech category did not improve the model fit. Thus, perceived dependence was not a significant predictor of friendship scores provided for speakers.

Finally, since prior research (e.g., Redford et al., 2018) has identified speech comprehensibility as a potentially important factor influencing people's affiliative interests in speakers with ASD, we were particularly interested in comprehensibility in the present study. Results from a mixed-effects model assessing differences in comprehensibility scores across speech conditions are shown in Table 67 and *post hoc* pairwise comparisons are shown in Table 68.

Table 67. Coefficient estimates from the mixed effects model predicting comprehensibility judgments from speech category (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
Intercept	2.37	0.10	24.86	< .001
L2	-0.90	0.13	-6.82	< .001
ASD	-0.38	0.13	-2.88	.005

Table 68. Pair-wise post hoc analysis assessing differences in comprehensibility scores between the speech categories

contrast	estimate	*SE*	*t*	*p*
L1 - L2	0.90	0.13	6.82	< 0.0001
L1 - ASD	0.38	0.13	2.88	0.01
L2 - ASD	-0.52	0.13	-3.94	< 0.0001

These analyses indicated that participants rated speakers with ASD as significantly less intelligible than L1 speakers but significantly more intelligible than L2 speakers (Table 68). Next, we assessed the extent to which comprehensibility scores predicted friendship scores across the three speech conditions. As was observed with physical competence, model comparisons indicated that a model predicting friendship scores from the main effects of speech category and comprehensibility (without an interaction) term, provided the best fit. However, since we had an *a priori* interest in the relative effect of comprehensibility on friendship scores across the speech categories, we also ran *post hoc* trend analysis assessing the relative effect of comprehensibility on friendship scores across the three speech categories. Results from a mixed-effects model assessing the main effects of speech category and comprehensibility are shown in Table 69 and *post hoc* trend analysis is provided in Table 70.

Table 69. Coefficient estimates from the mixed effects model predicting friendship scores from speech category and comprehensibility (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
(Intercept)	1.39	0.19	7.41	< .001
L2	0.11	0.13	0.81	.418
ASD	-0.04	0.12	-0.34	.736
Comprehensibility	0.26	0.07	3.86	< .001

Table 70. Post hoc trend analysis assessing the effect of comprehensibility on friendship scores for each speech category

Speech Category	Comprehensibility Trend	*SE*	lower.CL	upper.CL	*t*	*p*
L1	0.30	0.12	0.07	0.53	2.61	0.0098
L2	0.12	0.12	-0.13	0.36	0.94	0.35
ASD	0.34	0.11	0.12	0.56	3.02	0.003

Our analysis revealed that greater comprehensibility was significantly associated with greater friendship scores; however, this was primarily the case for L1 and ASD, but not L2 (Table 70). Thus, consistent with Redford et al., (2018), reduced comprehensibility scores were significantly associated with reduced friendship scores for ASD speakers (as well as L1 speakers).

Thus, to summarize the results of our analysis of speech-related predictors of friendship thus far: perceived physical competence predicted friendship scores but only for L1 speakers, perceived dependence was not associated with friendship scores provided for any of the speakers; and comprehensibility judgments predicted friendship scores for L1 and ASD (but not L2).

To further assess the relative contribution of each speech-related predictor to college students' friendship scores, we ran one model predicting friendship scores from speech category, comprehensibility, dependence, and competence. The results of this model are presented in Table 71.

Table 71. Coefficient estimates from the mixed effects model predicting friendship scores from audio type, comprehensibility, dependence and competence (Note: Neurotypical North American English speech – L1 – is the reference group)

	Estimate	*SE*	*t*	*p*
(Intercept)	1.26	0.25	5.02	< .001
L2	0.12	0.14	0.85	.399
ASD	-0.04	0.12	-0.31	.759
Comprehensibility	0.23	0.07	3.25	.001
Dependence	-0.05	0.07	-0.78	.437
Competence	0.13	0.07	1.83	.068

This model clarified that comprehensibility was the most robust (and in fact, only statistically significant) predictor of friendship scores provided to speakers. On the one hand, the fact that the link between reduced comprehensibility and lower friendship scores was not unique to speakers with ASD supports pragmatic accounts of negative evaluations of speakers with reduced comprehensibility (i.e., listeners might resent the extra work needed to achieve mutual understanding and blame/resent the interlocutor for this burden) in contrast to accounts that focus on perceptions of disability per se as a source of speech-related prejudice in the case of ASD (see Gertner et al., 1994; Perkins, 2007; Rice et al., 1991; and discussion by Redford et al., 2018). As one female participant (age 6) spontaneously exclaimed after one of the friendship prompts: "no I don't want to be friends with her – I can't understand what she's saying!"

That said, it is also worth noting complexity in relationships between comprehensibility and friendship scores. Namely, even though children rated L2 speech as significantly less intelligible than ASD speech, their friendship ratings for L2 and ASD speech did not differ statistically. Furthermore, our *post hoc* analysis indicated that comprehensibility was not significantly associated with friendship scores provided for L2 speakers, in contrast to what we would predict based on pragmatic accounts of speech-related prejudice. This suggests that comprehensibility was not the only driver of children's friendship scores; multiple factors likely operate regarding prejudicial responses to speech that vary along multiple dimensions beyond comprehensibility. However, for children in this study, comprehensibility was the only significant predictor of friendship scores provided for ASD speakers as neither perceived competence nor perceived dependence were associated with friendship scores provided for ASD speakers. All in all, within this study, reduced comprehensibility, rather than disability spread,

was associated with peer rejection against ASD speakers. Importantly, such peer rejection was not particularly evident in the present study (see Tables 67-68).

4.3.3.2.3. Predictors of Peer Rejection – Summary

In sum, for both the visual and speech stimuli, perceived physical competence predicted greater friendship scores, but only for typical images and L1 speech. Thus, perceptions of reduced physical competence detected in stimuli containing disability cues were not associated with reduced (or enhanced) friendship scores provided to these stimuli.

Interestingly, we observed a different pattern regarding the relationship between perceived dependence and friendship scores. Namely, for images, perceptions of greater dependence were associated with greater friendship scores for female (but not male) participants' ratings of the wheelchair images, but not for the other images. As well, perceptions of reduced competence were unrelated to friendship scores provided to speakers.

Finally, for speech, comprehensibility was the most robust predictor of friendship scores provided to speakers. Furthermore, (reduced) comprehensibility was the only predictor that was associated with reduced friendship scores provided to ASD speakers. Thus, disability spread (as assessed by the tendency to associate ASD speakers with reduced competence and enhanced dependence) was not associated with friendship scores provided for ASD speakers.

As noted, the fact that disability spread was generally not associated with peer rejection was also evident in children's evaluations of the images, since perceived physical competence was not associated with friendship scores for disability stimuli. Furthermore, perception of greater dependence was associated with greater friendship scores for female (but not male) participants' ratings of the wheelchair images, but not for the other images. Thus, for children,

spreading tendencies were generally not associated with peer rejection and, at least in the case of female participants, were associated with prosocial attitudes towards wheelchair images.

Taken together, the fact that disability spread was either unrelated or negatively related to peer rejection provides empirical support for theoretical depictions of disability spread as distinct from stigma (Liesener and Mills, 1999). In fact, at least in the case of female children evaluating wheelchair images, disability spread was associated with prosocial attitudes towards individuals with a disability, in contrast to the pattern observed with our college-aged sample (see general discussion). In the following chapter, we further discuss commonalities and differences in response patterns observed in the child and college student sample. While participants in these two samples responded in comparable ways for several dimensions, there were also clear differences. We now turn to discussion of these issues.

CHAPTER V

STUDY 1 GENERAL DISCUSSION

The fact that children and college students were evaluated in the same way, using identical stimuli and prompts supported the ability to compare response patterns observed in these two samples in a straightforward manner. This allowed us to explore commonalities as well as meaningful differences between these two demographics. Most notably, this analysis revealed significant differences in the expression of peer rejection as well as relationships between disability spreading and the expression of peer rejection.

5.1. Conceptual Differentiation

Children and college students responded to several measures in similar ways but also differed meaningfully on other measures. In terms of commonality, children and college students alike associated images depicting physical disabilities (but not foreignness) with low physical competence and high dependence. They also selectively associated foreign garb and foreign accent with being "from another country far away" but not with low competence or high dependence. Such findings reflect ways in which children and college students exhibited common inferential assumptions regarding the perceptual variability they evaluated.

On the other hand, there were also notable differences between college students and children. First, whereas children did not differ in their perceptions of dependence between the amputee and wheelchair images, college-aged participants provided significantly higher dependence scores for the wheelchair images in comparison to the amputee images. Thus, in contrast to children, college students likely attributed greater significance to the multiple disability cues in the wheelchair images (contracture and torticollis which may present due to a

range of factors) compared to an amputation when assessing the degree to which individuals' may "need help to do things".

Along those lines, children (but not college students) were somewhat unsure regarding whether the wheelchair images were "from America" or "another country far away". Thus, while children did associate the wheelchair images with high dependence and low physical competence, they may have lacked the conceptual knowledge to categorize this individual as "an American with a disability".

Thus, children may not have been familiar with individuals with disabilities such as severe cerebral palsy and thus may have assumed that the perceived differences were due to foreignness. This exposure account in supported by the fact that children whose parents indicated they'd had exposure to other children with a disability were more likely to associate wheelchair images with being "from America". Notably, there was also a trending age-related decrease in children's tendency to associate wheelchair images with being "from another country far away" and greater parental education exhibited a similar trend. Taken together, these findings suggest that, with age and increased knowledge, children gradually learn that disability status is uninformative regarding one's foreignness. On the other hand, the fact that, for the collapsed sample. the wheelchair images were less associated with being "from America" may be informative regarding early-emerging notions of ableism and nationalism.

5.2. Differences in Disability Spread Tendencies

Another important difference between children and college students related to the expression of disability spread. While there was clear evidence of disability spread in college students, the evidence was mixed for children. Namely, whereas college students associated autistic speakers with greater dependence and reduced physical competence in comparison to

speakers in the other two speech categories, children did not differ in their physical competence ratings across the speech conditions and showed only a trending tendency to associate ASD speakers with greater dependence. Thus, particularly for speech, disability spread was significantly more present in college students' responses in comparison to those of children.

5.3. Differences in the Expression of Peer Rejection

Yet another area of difference between the samples was evident in the friendship/peer rejection scores. First, consistent with previous research (Redford et al., 2018), college students provided significantly lower friendship scores for ASD speakers in comparison to the other speech categories. In contrast, children did not differ in their friendship ratings across the speech conditions, in contrast to the college sample in Redford et al., (2018). This age-related differential points to a possible window of opportunity to prevent age-related increases in prejudice toward ASD speakers.

Second, whereas comprehensibility was the only significant predictor of friendship scores provided for ASD speakers by children, perceived physical competence was the stronger predictor of friendship scores provided to speakers by college students. Nevertheless, while comprehensibility was not significantly associated with friendship scores provided to ASD speakers by college students, the higher comprehensibility scores that they provided to L1 speakers were related to the observed preference for L1 speakers over ASD speakers. Thus, comprehensibility scores were related to friendship scores provided to speakers by children and college students. These findings highlight the need for more research to continue exploring the role of comprehensibility and other factors in the expression of speech-related peer rejection and point to the possibility of developmental change in this regard.

Third, different response patterns between college students and children were also evident in the friendship ratings provided to the visual stimuli. Namely, whereas college students did not associate images depicting physical disabilities with lower friendship scores and showed a preference towards the amputee images, children provided lower friendship scores to the wheelchair and foreign images in comparison to the non-disabled images.

5.4. The Relationship Between Disability Spread and Peer Rejection

Finally, children and college students differed *vis-à-vis* relationships between spreading tendencies and expression of peer rejection. To recall, in our introduction we discussed two somewhat contrasting accounts regarding this relationship. On the one hand, associating reduced competence or enhanced dependence with an individual with a disability may sometimes stem from a prosocial desire to aid and support those perceived as in need. On this account, disability spread is conceptualized as distinct from stigma and prejudice (Liesener and Mills, 1999). On the other hand, associating reduced competence or enhanced dependence with disability cues may go together with the tendency to dehumanize individuals with a disability. There is some evidence from previous research to suggest that perceptions of dependence are associated with prejudice towards individuals with a disability (e.g., Parker, Monteith and South, 2020).

Along those lines, from the perspective of evolutionary and attachment approaches to human psychology, the positive association between perceptions of dependence and prosocial intentions is to be expected since these approaches view parental sensitivity to infant dependence as foundational to caregiving and social cognition more broadly (e.g., Hrdy, 2009). At the same time, evolutionary accounts are also consonant with aversive reactions to individuals perceived as dependent and low in competence, as affiliation with such individuals might compromise the group's fitness (e.g., Neuberg and Cottrell, 2013).

Interestingly, we found support for both accounts in our studies. Specifically, for children, disability spread was generally not associated with the expression of disability-related peer rejection. Thus, for children, neither perceived reduced physical competence nor perceived dependence were associated with friendship scores provided for ASD speakers or images depicting physical disabilities. In fact, for female children, perception of dependence was associated with enhanced interest in friendship. In contrast, for college students, there was a tendency for greater dependence scores to predict reduced friendship scores provided for autistic speakers as well as wheelchair images. An even more robust pattern was observed for physical competence and friendship ratings of speakers with ASD (though reduced physical competence predicted lower friendship scores regardless of speech condition). Overall, disability spread was associated with expression of peer rejection for college students, but not for children.

While future research is needed, one interpretation of these results *vis-à-vis* perceived competence is that such perceptions are likely "translated" to prosocial or antisocial intentions in relation to a range of contextual attributional factors at play (Weiner 2005;2006). Thus, humans likely consider the source of perceived competencies when incorporating such perceptions into their social preferences. As Weiner (2005; 2006) noted, the affiliative consequences of perceptions of reduced competence that are linked to uncontrollable factors (in this case: a physical disability) can differ drastically from perceptions of reduced competence that are linked to factors considered to be controllable (e.g., laziness).

Such attributional processes may explain why perceptions of reduced physical competence were not associated with friendship scores provided by college students and children for the images depicting physical disabilities. Since reduced physical competence was clearly attributable to a disability, it did not factor into participants' friendship preferences. On the other

hand, in the case of speakers with ASD, a salient attributional account for the cause of the perceived reduced physical competence may have been less readily accessible, thus explaining why college students may have "penalized" ASD speakers for their perceived reduced physical competence.

Beyond perceptions of physical competence, as noted, female children were the only group of participants for whom perceptions of dependence were associated with prosocial attitudes towards individuals with a physical disability. For the rest of the participants, perceptions of dependence were either not associated with friendship scores (male children) or associated with reduced friendship scores (college students). Thus, there is an important need to assess the conditions under which perceptions of dependence translate to prosocial intentions, factors that may foster this tendency, and the societal and interpersonal reasons why this tendency may subside in the case of perceptions of disability from childhood to emerging adulthood.

5.5. Summary

In study 1, we assessed college students (study 1a) and young children's (study 1b) tendency to distinguish between multimodal foreignness and disability cues, the inductive generalizations associated with such differentiation, and relations between these inductive generalizations and expression of peer rejection. First, in contrast to college students, children were unsure about whether ASD speakers and wheelchair images were "from America" or "another country far away". This suggests that children in the age range of those who participated in this study are still in the midst of constructing perceptual/conceptual differentiation between foreignness cues and disability cues. Second, whereas college students evaluated speakers with ASD in a similar manner to which they evaluated images of children

with physical disabilities (i.e., attributing high dependence and low physical competence), children generally did not seem to associate the variability they detected in the ASD speech with a latent disability concept. However, there was a trending tendency for children to associate (male) speakers with ASD with enhanced dependence, pointing to the need for future research in this regard. Third, whereas college students were biased against ASD speakers but showed a prosocial bias towards images depicting physical disabilities (particularly amputees), children were biased against wheelchair (and foreign) images but showed no bias in the case of speech. Finally, whereas disability spread was associated with peer rejection for college students, these tendencies were generally not associated with the expression of peer rejection for children and in one case were related to prosocial attitudes towards individuals in a wheelchair. These findings highlight the complex ways in which conceptual representations of the social world relate to the expression of peer rejection, and how such relationships may change developmentally.

In the following section, we raise several limitations with the way we operationalized disability spread in this first study. We then discuss how an alternative operationalization approach may shed further light on this phenomenon.

CHAPTER VI

AN IMPLICIT ASSOCIATIVE APPROACH TO ASSESSING FOREIGNNESS AND DISABILITY CONCEPTS

6.1. Overview

A key limitation of the approach taken in study 1 is that it relied on participants' ability to relate explicit lexical descriptors (e.g., being "from America", "good at running", "needing help to do things") to the stimuli. Since explicit talk about such concepts may be challenging for younger children, particularly in the case of speech processing (see Jones et al., 2016; Wagner, Greene-Havas, and Gillespie, 2010), this method may have limited our ability to accurately assess younger children's responses to the variability depicted in the stimuli.

Furthermore, regarding the dependence measure ("do they need a lot of help to do things"?) and the visual stimuli, under many circumstances, individuals with physical disabilities such as those depicted in our stimuli may in fact need help to achieve a range of tasks (e.g., climbing stairs). The extent to which judgments of perceived dependence should be regarded as constituting spreading – a term entailing perceptions of disability that are not directly related to an observed impairment – is therefore not clear. Taken together, these limitations indicated that further exploration was warranted of young children's intuitions regarding the social significance of variation in speech and appearance, using different methods.

To this end, in study 2, rather than explicitly gauging participants' impressions regarding the stimuli, we measured the extent to which participants would selectively associate speech produced by children with ASD with images depicting physical disabilities in a similar way to which they might associate foreign-accented speech with images depicting individuals dressed in foreign garb (Wagner, Clopper and Pate, 2014).

This approach facilitated more direct exploration of participants' tendency to "spread" a perceived impairment in one domain (i.e., speech) to another (i.e., physical disability) without relying upon explicit lexical judgments. Thus, the degree of selectivity in participants' speech-appearance associations is informative of the emergence of disability spread and its relation to inductive generalization in the case of foreignness concepts. Furthermore, comparing the extent to which participants selectively associate ASD speech with the wheelchair images (that depict a severe case of cerebral palsy, a centralized neurological disorder) to a greater extent than images depicting an amputee (a more peripheral physical disability) is also informative of the way factors such as severity and kind of disability relate to disability spreading tendencies.

To this end, in study 2 we assessed associations college students and children drew between speech and appearance, using the same stimuli from study 1. To recall, for speech, these included three categories: 1) neurotypical North American English-accented speech (L1), 2) American English-accented speech produced by children with autism spectrum disorder (ASD) and 3) neurotypical Spanish-accented English (L2). For visual stimuli, the stimuli included four illustration categories: children who are 1) wheelchair-bound with signs of contracture in the wrist and torticollis in the neck, 2) amputees, 3) foreign in their dress, or 4) non-disabled.

6.2. Research Questions

In this study, we examined the following research questions:

6.2.1. Will participants selectively associate disordered speech (ASD) with physical disability, but not with foreign garb?

6.2.2. Will participants selectively associate foreign-accented speech (L2) with foreign garb, but not with physical disability?

6.2.3. Attunement to disability features: will participants be more likely to associate disordered speech with wheelchair images than amputee images?

6.3. Predictions

The preregistration for this study along with detailed predictions is available at https://osf.io/bqut5/. In short, we predicted that college students would associate L1 speakers with typical images, L2 speakers with foreign images and ASD speakers with wheelchair images (more so than amputee images). In contrast, for children, whereas we did predict that children would selectively associate L1 speakers with typical images and L2 speakers with foreign images, we were unsure as to whether ASD speakers would be associated with images depicting physical disabilities.

6.4. Method

6.4.1. Participants

An *a priori* power analysis (using the "pwr" package in r) indicated that we would need 110 participants to detect a small effect size with 80% power. However, due to the exploratory nature of this study, we increased the sample to 136 (study 2a: 82 Female, 44 Male, 3 Gender Fluid, 1 Transgender, M_{age} = 19.9, SD = 1, age range = 18:26) and 140 (study 2b: 72 Female, 65 Male, M_{age} = 5;4, SD = 1.9, age range = 2:11) participants. For study 1a, college students participated in exchange for course credit at the University of Oregon. For study 1b, children were recruited within the Living Laboratory® at the Oregon Museum of Science and Industry (OMSI) in Portland, Oregon. Since the caregivers of 4 participants reported impossible ages for their children, their data were omitted from all non-age-related analyses. The final sample for study 1b was thus 141 for all non-age-related analyses (74 girls, 67 boys) and 137 (72 girls, 65 boys, M_{age} = 5.5, SD = 1.9, age range = 2:11) for all age-related analyses.

6.4.2. Materials

6.4.2.1. Speech Stimuli

See chapter 2, section 2.4.2

6.4.2.2. Illustration Stimuli

See chapter 2, section 2.4.2

6.4.3. Measures

6.4.3.1. Matching task – participants heard one of the three speech conditions (L1, L2, ASD) and were asked to select one of two illustrations which they believed best represented the child who produced the speech. The two illustrations presented on all trials always included a typical North American-presenting image and an image depicting the same individual from a different category (wheelchair, amputee, or foreign).

6.4.3.2. Reliability check – see chapter 2 section 2.4.3.

6.5. Procedure

All procedures followed the guidelines of the ethical advisory board at the University of Oregon. For study 2a, upon granting their consent, college students provided demographic information regarding their parent's levels of education as well as their age, gender, exposure to disability and mono/multilingualism. For study 2b, this information was provided by the caregivers. In both studies, the experiment was implemented/hosted on Gorilla (Gorilla.sc, Anwyl-Irvine et al., 2020). For study 2a, college students participated online without experimenter moderation. In contrast, for study 2b, testing took place on a one-to-one basis with a researcher according to Living Lab protocol on the OMSI museum floor.

All participants began with a brief comprehension task. Participants were first shown illustrations of a cow and lion side-by-side and were asked to listen carefully to the sound and

pick who is making the sound. They then heard a "moo" sound and were asked to make their selection. We then repeated the same procedure with side-by-side illustrations of a monkey and duck and an audio clip of monkey screeches. Upon completion of the comprehension task, participants completed the six test trials.

For all test trials, participants were presented with two illustrations side-by-side and were told to listen to a voice and select the child who is talking (Hirschfeld and Gelman, 1997; Wagner, Clopper and Pate, 2014). In each trial, one of the illustrations always depicted a typical North American child and the other depicted one of the three other illustration categories (wheelchair, amputee, foreign). Speech segments consisted of one of the three speech categories (Neurotypical-native English accent, L1; Neurotypical-Spanish accent, L2; speech of a child with autism spectrum disorder, ASD). Each child participated in six trials (two trials per speech condition). The order of the trials and position of the images were counter-balanced across participants. Finally, after the test trials, participants completed the reliability check task. The study took approximately 5 minutes to complete, and participants were compensated with research credits (in study 2a) or stickers (in study 2b) for completing the study.

6.6. Analysis

Our analysis plan was preregistered at https://osf.io/bqut5/. Mixed-effects analyses were performed in the R statistical language (Version 4.1.3; R Core Team. 2020) using the lme4 package (Bates et al., 2015), and *p*-values of the models were calculated using the lmerTest package (Kuznetsova et al., 2017). In these models, categorical variables were contrast coded such that the intercept of a model represents the mean of all data points in the data set. All continuous predictors were standardized and mean-centered to facilitate the interpretation of

intercepts. *Post-hoc* analyses were conducted through pairwise comparisons using the emmeans package (Lenth, 2022).

For all trials, selections of the typical child were coded as 1 and all other selections were coded as 0. We then ran a mixed effects logistic regression model predicting selection (1 = typical, 0 = impairment/foreign) from the main effects and interaction between speech condition (L1, L2, ASD) and visual comparison type (Typical vs. Foreign, Typical vs. Amputee, Typical vs. Wheelchair).

Alpha was set at .05 when examining the relationship between variables but we discuss the overall support for a given hypothesis based on the pattern of findings observed across analyses. Finally, we used the Tukey method to correct for multiple comparisons in our *post hoc* analyses.

CHAPTER V11

STUDY 2a (COLLEGE-AGED SAMPLE) RESULTS AND DISCUSSION

In study 2a, we assessed college students' associations between speech variability (three categories: L1-neurotypical North American English, L2-neurotypical Spanish English, and ASD - North American English produced by children with ASD) and appearance (Two categories with a disability: wheelchair, amputee; and two categories without disability: North American, foreign). The overarching goal of our analysis was to assess the extent to which participants would selectively associate disordered speech with physical disability in a similar manner to which they might associate foreign-ccented speech with foreign garb. As noted, we predicted that college students would associate L1 speakers with typical images, L2 speakers with foreign images and ASD speakers with wheelchair images (more so than amputee images).

Descriptive statistics for parental education and participants' age, gender, exposure to disability, multilingualism, race, and ethnicity can be found in the supplementary materials (section 2).

7.1. Data Cleaning and Reliability Check Assessment

Of the 136 participants assessed, 88 (64.7%) indicated that it was not easy for them to understand Hebrew. In contrast, 16 (11.7%) indicated that Hebrew was very easy to understand and 32 (23.5%) indicated that Hebrew was a little easy to understand. Since this study was administered online, the reliability check served as a useful tool to filter out participants who may not have been paying adequate attention to the study. Thus, we ran our analysis on the 88 participants who indicated it was not easy to understand Hebrew (55 female, 29 male, 2 gender

fluid, 2 transgender, M_{age} = 19.9, SD = 2, range = 18:26 years). Importantly, the general pattern of results was consistent for the aggregate sample as well (see Figure 18).

7.2.Research questions

To assess our research questions, we ran a mixed effects logistic regression model predicting participants' binary choice (1 = typical image selection, 0 = other image selection) from visual comparison (typical vs. foreign; typical vs. amputee; typical vs. wheelchair), speech category (L1; L2; ASD), and the interaction between the two with a random intercept for participant. The results of this model are presented below. In this model, we set L1 as the reference group for speech and typical versus foreign as the reference group for the visual comparison variable. Thus, coefficients in the regression table should be interpreted as representing deviations from the likelihood of typical image selection observed for participants listening to L1 in typical versus foreign visual comparison trials. For example, a negative coefficient for L2 would indicate that participants listening to L2 in typical versus foreign visual-comparison trials were significantly more likely to choose foreign images than participants listening to L1 in these trials. Below we provide descriptive statistics with binomial tests assessing deviance from chance, the results of the mixed-effects regression model, *post hoc* tests and a visualization are provided, followed by a narrative explanation of the findings in light of the research questions.

Table 72. Descriptive Statistics and Binomial tests of participant's binary choice (typical vs. other) in relation to the speech category and visual comparison.

Speech Category	Visual Comparison	% Typical Selection	% Other Selection	Two-tailed Binomial Significance Test
L1	Typical/Foreign	85.4%	14.5%	0.00000001
L1	Typical/Amputee	73.5%	26.4%	0.0008
L1	Typical/Wheelchair	81.9%	18%	0.0000004
L2	Typical/Foreign	47.1%	52.8%	0.7838463
L2	Typical/Amputee	60.6%	39.3%	0.1237314
L2	Typical/Wheelchair	54.8%	45.2%	0.5257733
ASD	Typical/Foreign	75.4%	24.6%	0.00008
ASD	Typical/Amputee	74.2%	25.8%	0.00017
ASD	Typical/Wheelchair	30.2%	69.8%	0.0055

Table 73. Coefficient estimates from the mixed effects model predicting image category selection (typical = 1, other = 0) from the main effects and interactions between visual comparison and speech category (Note: North American-presenting images without disability vs. foreign presenting images without disability are the visual comparison reference group and neurotypical North American English speech – L1 – is the reference group for speech).

	Estimate	*SE*	*z*	*p*
Intercept	1.8222	0.37129	4.90795	0.0000009
Typical vs. Amputee	-0.7838	0.48476	-1.6169	0.10589
Typical vs. Wheelchair	-0.265	0.49587	-0.5357	0.59212
L2	-1.9518	0.46796	-4.1708	0.00003
ASD	-0.66693	0.47302	-1.4099	0.15856
Typical vs. Amputee x L2	1.36323	0.62774	2.1716	0.02988
Typical vs. Wheelchair x L2	0.60213	0.62943	0.9566	0.33875
Typical vs. Amputee x ASD	0.72055	0.64367	1.1194	0.26294
Typical vs. Wheelchair x ASD	-1.7636	0.66220	-2.663	0.00774

Table 74. Post hoc simple contrasts analysis assessing differences in likelihood of typical selection between visual comparison groups for each speech condition.

contrast	audio	estimate	*SE*	*z*	*p*
(Typical/Foreign) - (Typical/Amputee)	L1	0.7838	0.4847	1.6169	0.23841
(Typical/Foreign) - (Typical/Wheelchair)	L1	0.26567	0.4958	0.5357	0.8537
(Typical/Amputee) - (Typical/Wheelchair)	L1	-0.5181	0.4641	-1.1164	0.5038
(Typical/Foreign) - (Typical/Amputee)	L2	-0.57941	0.3904	-1.4839	0.29863
(Typical/Foreign) - (Typical/Wheelchair)	L2	-0.33646	0.3850	-0.8737	0.6569
(Typical/Amputee) - (Typical/Wheelchair)	L2	0.2429	0.3731	0.6510	0.7917
(Typical/Foreign) - (Typical/Amputee)	ASD	0.0632	0.4216	0.1500	0.98766
(Typical/Foreign) - (Typical/Wheelchair)	ASD	2.0292	0.4389	4.6233	0.000011
(Typical/Amputee) - (Typical/Wheelchair)	ASD	1.966	0.4342	4.5270	0.00001

Table 75. Post hoc simple contrasts analysis assessing differences in likelihood of typical selection between speech conditions in relation to each visual comparison group.

contrast	*Visual comparison*	*estimate*	*SE*	*z*	*p*
L1 - L2	*Typical/Foreign*	*1.951813*	*0.467963*	*4.17086*	*0.000089*
L1 - ASD	*Typical/Foreign*	*0.666930*	*0.473027*	*1.4099*	*0.33571*
L2 - ASD	*Typical/Foreign*	*-1.284883*	*0.41804*	*-3.0735*	*0.00599*
L1 - L2	*Typical/Amputee*	*0.588576*	*0.414581*	*1.41968*	*0.33068*
L1 - ASD	*Typical/Amputee*	*-0.05362*	*0.43365*	*-0.12365*	*0.99160*
L2 - ASD	*Typical/Amputee*	*-0.64220*	*0.39856*	*-1.61127*	*0.24080*
L1 - L2	*Typical/Wheelchair*	*1.349678*	*0.4277*	*3.15566*	*0.004563*
L1 - ASD	*Typical/Wheelchair*	*2.430539*	*0.465967*	*5.21611*	*0.0000005*

L2 - ASD	*Typical/Wheelchair*	*1.080861*	*0.405738*	*2.66393*	*0.02108*

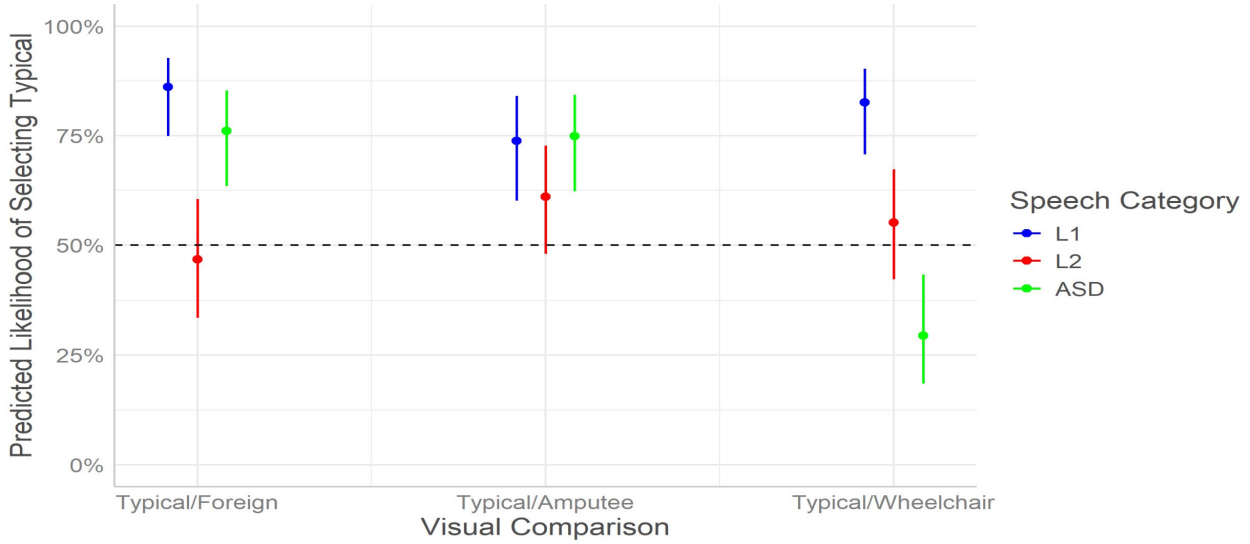

Figure 17. **Predicted likelihood of selecting Typical images in relation to visual comparison and speech category (filtered sample). Points represent estimates from the mixed effects regression model and lines represent 95% confidence intervals based on the standard errors.**

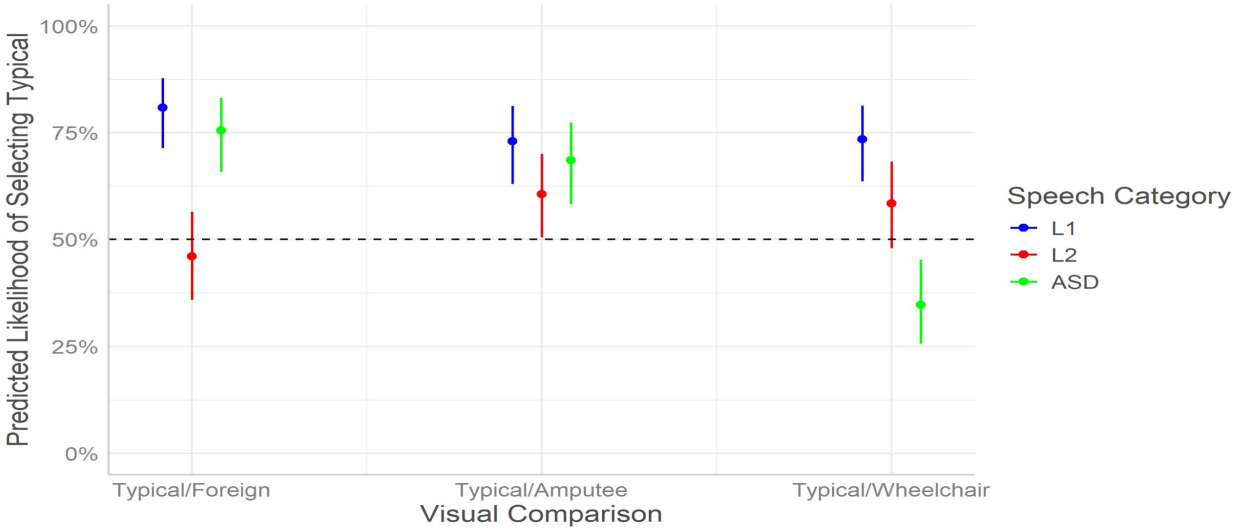

Figure 18. **Predicted likelihood of selecting Typical images in relation to visual comparison and speech category (unfiltered sample). Points represent estimates from the mixed effects regression model and lines represent 95% confidence intervals based on the standard errors.**

7.2.1. Will participants selectively associate disordered speech with physical disability but not with foreign garb?

When listening to ASD speakers, participants were significantly more likely to select wheelchair images in typical versus wheelchair visual-comparison trials than they were to select foreign images in typical versus foreign visual-comparison trials or amputee images in typical versus amputee visual-comparison trials (see table 74 and Figure 17). Furthermore, in typical versus wheelchair visual-comparison trials, participants were significantly more likely to select wheelchair images when listening to ASD speech in comparison to when listening to L1 or L2 speech (see table 75 and Figure 17). Finally, in typical versus foreign visual-comparison trials, participants were significantly more likely to select typical images when listening to ASD speech than when listening to L2 speech (Table 75 and Figure 17).

Taken together, these results indicate that participants were clearly inclined to selectively associate ASD speech with wheelchair images but not with foreign garb, as predicted.

7.2.2. Will participants selectively associate foreign-accented speech with foreign garb but not with physical disability?

For typical versus foreign visual-comparison trials, participants were significantly more likely to select foreign images when listening to L2 speech in comparison to both L1 and ASD speech (Table 75 and Figure 17). However, participants were not inclined to associate L2 speech with foreign garb at levels above chance (see table 72 and Figure 17). Thus, we observed a relative (i.e., in comparison to L1 and ASD) rather than absolute effect vis-à-vis L2/foreign-garb mapping tendencies.

As well, for typical versus wheelchair visual-comparison trials, participants were significantly more likely to select wheelchair images when listening to L2 speech in comparison to L1 speech, but significantly less likely to select wheelchair images when listening to L2 speech in comparison to when listening to ASD speech (see Table 75 and Figure 17).

Thus, participants associated L2 speech with foreign garb to a significantly greater degree than they associated L1 or ASD speech with foreign garb, as predicted. On the other hand, somewhat unexpectedly, they were also more likely to associate L2 speech with wheelchair images in comparison to L1 speech. Importantly, despite the significant difference between L1 and L2 vis-à-vis wheelchair selection, L2 speech was paired with typical images more frequently (55%) than with wheelchair images (45%) in typical versus wheelchair visual-comparison trials (though the tendency to select typical images in these cases was not significantly different from chance – see table 72).

In sum, the associations between ASD and wheelchair images were similar to the associations between L2 and foreign garb, as predicted. Surprisingly (and strikingly), the ASD-wheelchair association was perhaps even more robust than the L2-foreign garb association, since the former exhibited an absolute effect (e.g., significantly different from chance) while the latter only a relative effect (i.e., significantly different than the other speech categories – see Table 72 and Figure 17).

7.2.3. Attunement to disability features: will participants be more likely to associate disordered speech with wheelchair images than amputee images?

When listening to ASD speakers, participants were significantly more likely to select wheelchair images in typical versus wheelchair visual-comparison trials than they were to

select amputees in typical versus amputee trials (Table 74 and Figure 17). Furthermore, for typical versus amputee visual-comparison trials, participants were no more likely to select amputee images when listening to ASD speakers in comparison to when listening to L1 or L2 speakers (Table 75 and Figure 17). Thus, the tendency to associate ASD speakers with a physical disability was evident only in the case of wheelchair images (not amputee images), as predicted. In fact, no meaningful differences in response patterns across speech categories were observed for typical versus amputee visual-comparison trials and participants tended to associate all three speech categories with typical images in these comparisons. Thus, no speech category evoked associations with amputee images, in contrast to the other speech-image associations described above.

The fact that participants selectively associated ASD speech with the wheelchair illustrations but not the amputee illustrations (as predicted) is noteworthy. Namely, participants appeared to associate disordered speech with images that contained multiple disability cues (i.e., wheelchair with signs of contracture in the wrist and torticollis in the neck) but not images that contained an isolated/peripheral disability cue (i.e., amputation). In fact, in stark contrast to the wheelchair images, amputee images did not selectively evoke associations with any of the three speech categories. Thus, participants appear to have viewed an amputation as irrelevant to speech and thus as less inductively rich than the foreign or wheelchair images. The emerging conclusion from the data is that, rather than occurring "promiscuously", spreading tendencies appear to operate in relation to the perceived severity and nature of a disability cue.

7.3. Summary

In sum, participants were clearly inclined to selectively associate ASD speech with wheelchair images but not with foreign garb or amputees. As well, participants reliably associated L2 speech with foreign garb to a significantly greater degree than they associated L1 or ASD speech with foreign garb. Taken together, these findings indicate that participants tended to "spread" a perceived impairment detected in one domain (speech) to another domain (physical disability) in a similar manner to which they generalized perceived variability in speech (i.e., an accent) to appearance (i.e., foreign garb).

The results from the college-aged sample are also consistent with those we observed in study 1a. To recall, in study 1a, college students associated ASD speakers with reduced physical competence and enhanced dependence, in a similar manner to which they evaluated images depicting physical disabilities. Complementing these findings, in study 2a, college students associated ASD speakers with those same images depicting physical disabilities, in a similar manner to which they associated L2 speakers with images depicting foreign garb. Strikingly, the tendency to associate ASD speakers with wheelchair images was more robust than the tendency to associate L2 speakers with foreign garb.

We now turn to our investigation of children's associations between speech variability and appearance.

CHAPTER VIII

STUDY 2b (CHILD SAMPLE) RESULTS AND DISCUSSION

In study 2b, we measured the associations children (Ages 2.5-11 years) draw between speech variability and appearance. To recall, based on previous literature (Hirschfeld and Gelman, 1997; Wagner, Copper and Pate, 2014) we predicted that children would associate L2 speech with foreign garb. However, previous literature has never assessed whether children would associate disordered speech with physical disabilities, and we did not pre-register a prediction regarding this issue. The key question motivating this study was, therefore, assessing the extent to which children would selectively associate speech produced by children with ASD with images depicting physical disabilities in a similar manner to which they might be expected to selectively associate L2 speech with foreign garb. The results are thus informative not only regarding the developmental emergence of disability spread, but also regarding the degree of conceptual differentiation children draw between foreignness and disability.

Descriptive statistics for parental education and children's age, gender, exposure to disability, multilingualism, race, and ethnicity can be found in the supplementary materials (section 2).

8.1. Data Cleaning and Reliability Check Assessment

Of the 141 participants assessed, 72 (51%) indicated that it was not easy for them to understand Hebrew. In contrast, 40 (28.3%) indicated that Hebrew was very easy to understand and 29 (20.5%) indicated that Hebrew was a little easy to understand. As in Study 1b, the correlation between age and these comprehensibility responses was sizable (r = -.42, 95% CI: -0.55, -0.28, $p < 0.0001$).

To recall, part of the motivation to assess children's inductive generalizations with the matching task was that our initial explicit approach relied on children's ability to express their evaluations of speakers via lexically designated dimensions such as whether the speakers were "easy to understand", "good at running", "needed help to do things", or were "from America or another country far away". Since this approach may have been particularly challenging for young children, it may have interfered with our ability to assess children's intuitions regarding relationships between speech variability and appearance. Thus, we reasoned that a simpler task in which children simply must point to one of two potential candidate images when hearing a speaker would potentially shed light on younger children's evaluations of speakers. Due to our interest in capturing younger children's intuitions with the matching task, we were interested in assessing participants' performance irrespective of their ability to pass the reliability check. Thus, whereas for the explicit approach we filtered out participants who failed the reliability check, here we begin with reporting children's performance across the whole sample. However, to maintain consistency across studies and to explore potential differences between participants who passed and failed the reliability check, we assessed reliability check as a moderator in our analysis of the aggregate sample and also ran our analysis only on the 72 participants (45 girls, 26 boys, M_{age} = 6.2, SD = 1.9, range = 2:11) who passed the reliability check (i.e., reported difficulty understanding Hebrew, a language with which they were unfamiliar).

8.2.Research questions

First, for the aggregate sample, likelihood ratio tests indicated that accounting for age or reliability check in a model predicting participants' binary choice (1 = typical image selection, 0 = other image selection) from visual comparison (typical vs. foreign; typical vs. amputee; typical vs. wheelchair) and speech category (L1; L2; ASD), did not improve the model fit. Therefore, to

assess our research questions, we ran a mixed effects logistic regression model predicting participants' binary choice (1 = typical image selection, 0 = other image selection) from visual comparison (typical vs. foreign; typical vs. amputee; typical vs. wheelchair) and speech category (L1; L2; ASD), with a random intercept for participants without accounting for participants' age or reliability check scores.

Below we provide descriptive statistics with binomial tests assessing deviance from chance, the results of the mixed effects regression model described above, *post hoc* tests and a visualization, followed by a narrative explanation of the findings in light of the research questions.

Table 76. Descriptive statistics and binomial tests of participant's binary choice (typical vs. other) in relation to the speech category and visual comparison.

audio	visa	% Typical Selection	% Other Selection	Two-tailed Binomial Significance Test
L1	Typical/Foreign	63.4%	36.6%	0.0124
L1	Typical/Amputee	54.3%	45.6%	0.465
L1	Typical/Wheelchair	52.8%	47.2%	0.671
L2	Typical/Foreign	39.1%	60.8%	0.047
L2	Typical/Amputee	52.8%	47.2%	0.671
L2	Typical/Wheelchair	45.1%	54.8%	0.4
ASD	Typical/Foreign	57.3%	42.7%	0.2
ASD	Typical/Amputee	54.9%	45.1%	0.4
ASD	Typical/Wheelchair	54.4%	45.6%	0.465

Table 77. Coefficient estimates from the mixed effects model predicting image category selection (typical = 1, other = 0) from the main effects and interactions between visual comparison and speech category (Note: North American presenting images without disability versus foreign presenting images with disability are the visual comparison reference group and neurotypical North American English Speech is the reference group for speech).

	Estimate	*SE*	*z*	*p*
(Intercept)	0.59387	0.2329	2.5492	0.01079
Typical vs. Amputee	-0.3677	0.3174	-1.1582	0.24676
Typical vs. Wheelchair	-0.4992	0.3211	-1.55445	0.12007
L2	-1.039	0.3209	-3.2369	0.00120
ASD	-0.3003	0.322	-0.93244	0.35110
Typical vs. Amputee x L2	0.9075	0.4534	2.00157	0.0453
Typical vs. Wheelchair x L2	0.7243	0.45328	1.5979	0.11006
Typical vs. Amputee x ASD	0.2788	0.45189	0.6170	0.53722
Typical vs. Wheelchair x ASD	0.4318	0.45643	0.94606	0.3441

Table 78. Post hoc simple contrasts analysis assessing differences in likelihood of typical selection between visual comparison groups for each speech condition.

contrast	*audio*	*estimate*	*SE*	*z*	*p*
(Typical/Foreign) - (Typical/Amputee)	*L1*	*0.36771*	*0.31748*	*1.1582*	*0.47827*
(Typical/Foreign) - (Typical/Wheelchair)	*L1*	*0.49921*	*0.32114*	*1.5544*	*0.26564*
(Typical/Amputee) - (Typical/Wheelchair)	*L1*	*0.13149*	*0.31711*	*0.4146*	*0.9095*
(Typical/Foreign) - (Typical/Amputee)	*L2*	*-0.53981*	*0.31932*	*-1.69*	*0.2087*
(Typical/Foreign) - (Typical/Wheelchair)	*L2*	*-0.22509*	*0.315*	*-0.7124*	*0.75615*
(Typical/Amputee) - (Typical/Wheelchair)	*L2*	*0.31471*	*0.315*	*0.9961*	*0.5792*
(Typical/Foreign) - (Typical/Amputee)	*ASD*	*0.08888*	*0.3171*	*0.2802*	*0.9576*
(Typical/Foreign) - (Typical/Wheelchair)	*ASD*	*0.06739*	*0.318*	*0.2118*	*0.9755*
(Typical/Amputee) - (Typical/Wheelchair)	*ASD*	*-0.021497*	*0.313*	*-0.0686*	*0.9974*

Table 79. Post hoc simple contrasts analysis assessing differences in likelihood of typical selection between speech conditions for each visual comparison group.

contrast	*Visual Comparison*	*estimate*	*SE*	*z*	*p*
L1 - L2	*Typical/Foreign*	*1.039*	*0.32*	*3.2369*	*0.0034*
L1 - ASD	*Typical/Foreign*	*0.30032*	*0.322*	*0.9324*	*0.6197*
L2 - ASD	*Typical/Foreign*	*-0.73869*	*0.3206*	*-2.303*	*0.0552*
L1 - L2	*Typical/Amputee*	*0.131491*	*0.3171*	*0.4146*	*0.9095*
L1 - ASD	*Typical/Amputee*	*0.02149*	*0.3131*	*0.0686*	*0.997*
L2 - ASD	*Typical/Amputee*	*-0.109*	*0.3160*	*-0.3480*	*0.9354*
L1 - L2	*Typical/Wheelchair*	*0.3147*	*0.315917*	*0.9961*	*0.5792*
L1 - ASD	*Typical/Wheelchair*	*-0.131*	*0.31711*	*-0.414*	*0.909*
L2 - ASD	*Typical/Wheelchair*	*-0.4462*	*0.3139*	*-1.421*	*0.329*

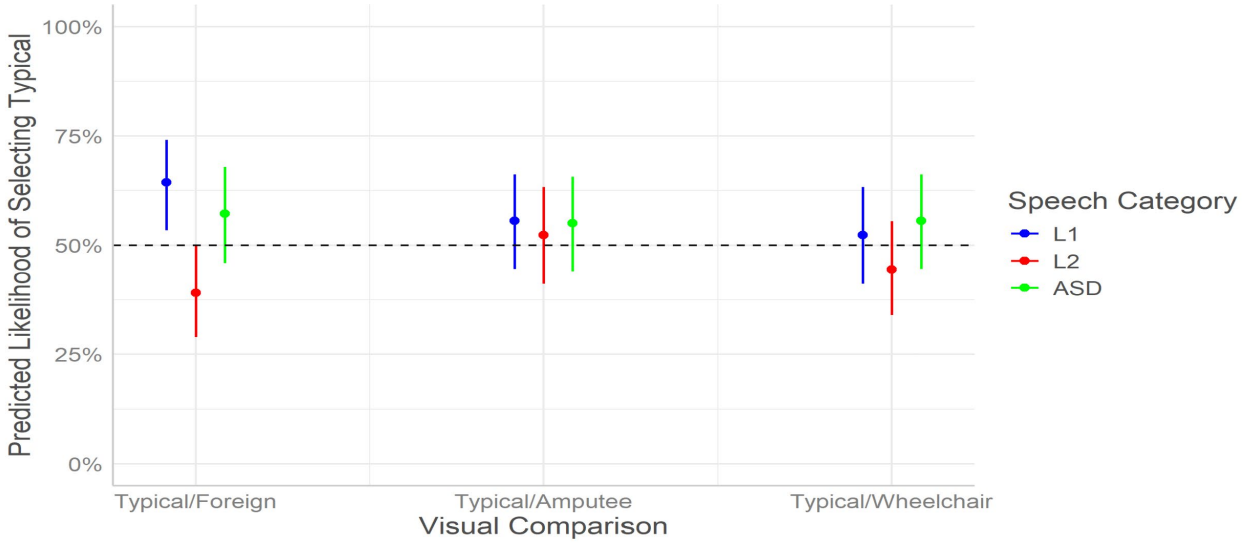

Figure 19. **Predicted likelihood of selecting Typical images in relation to visual comparison and speech category for all children (even those who did not pass the reliability check). Points represent estimates from the regression model and lines represent 95% confidence intervals based on the standard errors.**

For the aggregate sample, for typical vs. foreign visual-comparison trials, children were significantly more likely to select typical images when listening to L1 speech in comparison to L2 speech, but not in comparison to ASD speech (Table 79 and Figure 19). Put another way, for typical vs. foreign visual-comparison trials, children were significantly more likely to select foreign images when listening to L2 speech in comparison to L1

speech, as predicted, and to a lesser (non-significant) extent, in comparison to ASD speech as well.

Notably, beyond the Typical vs. Foreign comparison, there were no significant associations between speech categories and image selection for the other two comparison trials (Typical vs. Amputee and Typical vs. Wheelchair). Thus, for the aggregate sample, when faced with typical vs. foreign comparisons, children reliably associated L1 and ASD speech with typical images and L2 speech with foreign images. In contrast, there was no evidence that children associated ASD speech with a physical disability in the way they associated L2 speech with foreign garb and there was no age-related change in this regard.

As noted, accounting for the reliability check did not improve our model fit. Illustrative of this, in Figure 20 we provide a visualization of the model from only participants who passed the reliability check. As is apparent, this visualization looks much the same as the visualization for the aggregate sample.

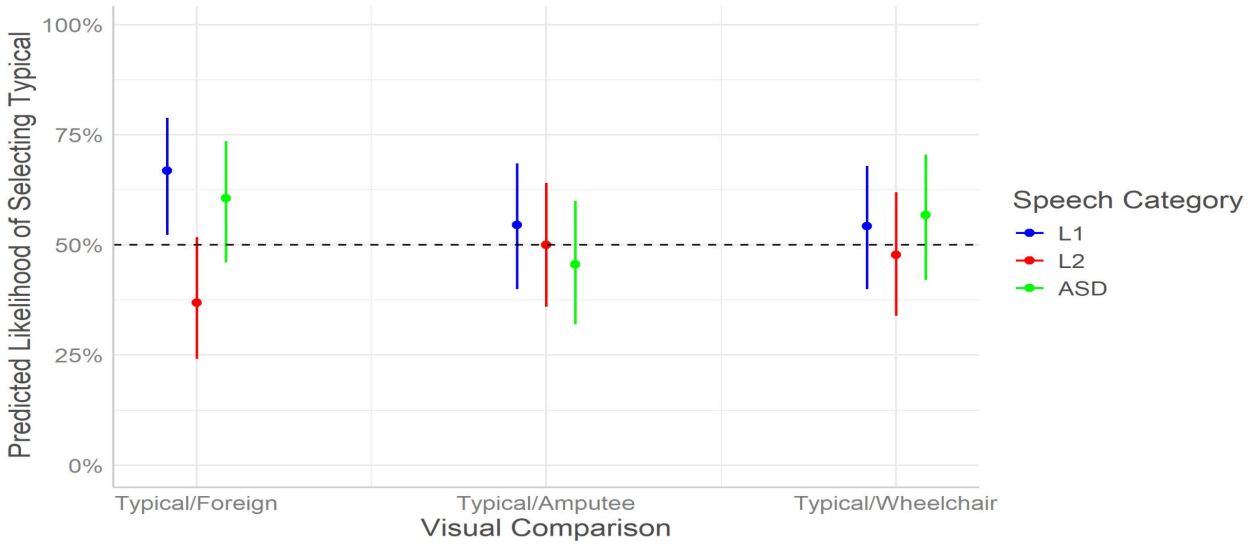

Figure 20. **Predicted likelihood of selecting Typical images in relation to visual comparison and speech category for children who passed the reliability check. Points represent estimates from the regression model and lines represent 95% confidence intervals based on the standard errors.**

Thus, to address our research questions for the aggregate data, children selectively associated foreign-accented speech with foreign garb but did not selectively associate disordered speech with physical disability. As well, children were no more likely to associate disordered speech with wheelchair images than amputee images and there was no measurable age-related change in this regard.

Beyond replicating and extending previous work (Hirschfeld and Gelman, 1997; Wagner, Copper and Pate, 2014), children's selective tendency to associate L2 speech with foreign images was potentially informative of long-lasting debates between "rich" conceptual and "low-level" perceptual accounts of children's representation of relationships between speech variability and social differences. Namely, "rich" conceptual accounts suggest that children's representations of social groups entail expectations of group-defining ways of speaking (i.e., accents) with corresponding ways of dressing. On this account, the tendency to map foreign-sounding speech onto foreign dress reflects specific expectations that these two categories are conceptually linked (e.g., Chomsky, 1987; Hirschfeld and Gelman, 1997). In contrast, "low level" perceptual accounts argue that children simply associate less familiar/typical auditory stimuli with less familiar/typical visual stimuli. Thus, children may associate foreign-sounding speech with foreign-looking dress due to shared atypicality of this variability rather than conceptual understanding that these sources of variability are indices of social group affiliation (e.g., Ramsey, 1987).

To recall, children in the present study selectively mapped foreign garb, rather than amputee or wheelchair images, to L2 speech. Furthermore, they selectively mapped foreign garb onto L2 speech, but not ASD speech. Taken together, these results suggest that participants did not merely map atypical speech to atypical appearance, as low-level perceptual accounts would

suggest. Rather, children's associative patterns imply that their foreignness concepts entail expectations that specific kinds of speech variability (i.e., an accent) but not others (i.e., a speech disorder) relate to specific kinds of appearance variability (i.e., garb) but not others (i.e., physical disability). This degree of selectivity appears to be more consistent with rich conceptual accounts of children's understanding of society. However, children's tendency to selectively map foreign-accented speech to foreign garb could still be attributed to statistical learning processes rather than a conceptual understanding of "the importance of cultural variation per se in a relatively broad sense" (Kuczaj and Harbaugh, 1982:225). More research is therefore needed to adjudicate between these competing accounts.

In contrast to the associative patterns related to children's foreignness concepts evident for the aggregate sample, distinct associative patterns pertaining to children's disability concepts and related disability spread tendencies were not apparent in the aggregate sample. Specifically, children showed no measurable tendency to associate ASD speech with images depicting physical disabilities and no age-related change in this regard. The fact that children displayed foreignness generalizations but not disability spread in the aggregate analysis suggests that the detection of foreignness in speech and the inductive generalizations associated with this detection may precede analogous processes occurring in regard to perception of a speech disorder. However, since the L2 and ASD speech samples were not matched in terms of degree of foreignness/disability, strong conclusions regarding the relative primacy or salience of one sort of variability over the other are not warranted.

8.3.Exploratory Analysis

To further probe children's associative tendencies vis-à-vis ASD speakers and physical disabilities, we explored the extent to which accounting for speaker gender would add explanatory value. However, a likelihood-ratio test indicated that accounting for speaker gender in a model predicting participants' binary choice (1 = typical image selection, 0 = other image selection) from visual comparison (typical vs. foreign; typical vs. amputee; typical vs. wheelchair) and speech category (L1; L2; ASD), did not improve the model fit (in a similar manner to which accounting for age or reliability check did not improve the model fit). However, further exploratory analysis indicated that accounting simultaneously for both age and speaker gender did significantly improve the model fit. Namely, for typical vs. wheelchair visual-comparison trials, there was a trending age-related increase in the tendency to associate male (but not female) ASD speakers with wheelchair images. In contrast, there was no such tendency for typical vs. amputee images. As well, for typical vs. foreign visual comparison trials, there was a significant age-related increase in the tendency to associate male (but not female) ASD speakers with typical images (see Table 80).

Table 80. Post hoc trend analysis assessing age related change in likelihood of typical selection in relation to the visual comparison group, speech category and speaker gender.

Visual Comparison	Audio	Speaker Gender	Age Trend	*SE*	Asymp. LCL	asymp. UCL	*z*	*p*
typ/for	L1	female	1.3238	0.5016	0.340	2.307	2.63	0.00831
typ/amp	L1	female	0.6446	0.3925	-0.124	1.4140	1.64	0.10052
typ/wc	L1	female	0.5045	0.38631	-0.252	1.2616	1.305	0.1915
typ/for	L2	female	0.7579	0.420	-0.0667	1.5827	1.8013	0.0716
typ/amp	L2	female	-0.541	0.3425	-1.2131	0.1296	-1.581	0.1137

typ/wc	L2	female	-0.307	0.36495	-1.022	0.4079	-0.842	0.3997
typ/for	ASD	female	-0.145	0.2747	-0.6837	0.3931	-0.528	0.5969
typ/amp	ASD	female	0.0273	0.33083	-0.6210	0.6757	0.0826	0.9341
typ/wc	ASD	female	-0.064	0.2999	-0.6525	0.5231	-0.215	0.8291
typ/for	L1	male	0.097	0.33428	-0.5581	0.7522	0.2903	0.77156
typ/amp	L1	male	-0.785	0.34346	-1.4587	-0.1124	-2.287	0.0221
typ/wc	L1	male	-0.212	0.2966	-0.7937	0.3689	-0.716	0.47388
typ/for	L2	male	-0.723	0.33412	-1.3785	-0.0688	-2.166	0.03031
typ/amp	L2	male	0.0502	0.3364	-0.6092	0.7097	0.1492	0.8813
typ/wc	L2	male	0.6925	0.401	-0.0934	1.4785	1.7269	0.0841
typ/for	ASD	male	1.362	0.5572	0.27	2.455	2.446	0.01444
typ/amp	ASD	male	-0.268	0.35636	-0.967	0.4298	-0.753	0.45105
typ/wc	ASD	male	-0.664	0.37112	-1.3915	0.0632	-1.789	0.0735

To further probe the nature of age-related increase in the tendency to associate male ASD speakers with wheelchair images, we added reliability check as a categorical moderator (pass/fail) to the logistic regression model. This revealed that, for participants who passed the reliability check, the age-related increase in the tendency to associate male (but not female) ASD speakers with wheelchair images was significant (see Table 81). In contrast, there was no significant age-related change in the tendency to associate ASD speakers with amputee images in typical vs. amputee visual comparison trials.

Thus, particularly for children who passed the reliability check, there was an age-related increase in the tendency to selectively associate male (but not female) ASD speakers with

wheelchair images, but not with foreign or amputee images. This provides further evidence for conceptual differentiation (as ASD speakers were not associated with foreign garb), disability spread (since ASD speakers were associated with wheelchair images) and attunement to disability features (since ASD speakers were associated with wheelchair images but not amputee images).

At the same time, we also observed several anomalous age-related patterns which were difficult to interpret. Specifically, we observed trending age-related increases in the tendency to associate male L1 speakers and female L2 speakers with amputee images (see Table 81). Furthermore, there was a significant age-related increase in the tendency to associate female L2 speakers with typical images in typical vs. foreign visual comparison trials. Thus, participants' associations vis-à-vis male L1 and female L2 speakers did not consistently match the expected patterns (i.e., female L2 speakers were not associated with foreign garb; female L2 and male L1 speakers were associated with an amputation).

Table 81. Post hoc trend analysis assessing age related change in likelihood of typical selection in relation to the visual comparison group, speech category, speaker gender and participants reliability check.

Visual comparison	Speech Category	Reliability Check	speaker gender	Age trend	*SE*	*z*	*p*
typ/for	L1	fail	female	0.771	0.572	1.3477	0.1777
typ/amp	L1	fail	female	0.009	0.6422	0.0144	0.9884
typ/wc	L1	fail	female	0.169	0.674	0.2515	0.8013
typ/for	L2	fail	female	-0.59	0.8827	-0.674	0.4999
typ/amp	L2	fail	female	0.588	0.7519	0.7824	0.4339
typ/wc	L2	fail	female	-1.15	1.0485	-1.099	0.2715

typ/for	ASD	fail	female	0.021	0.7521	0.0281	0.9775
typ/amp	ASD	fail	female	0.404	0.6212	0.6505	0.5153
typ/wc	ASD	fail	female	-0.12	0.5642	-0.227	0.8204
typ/for	L1	pass	female	2.489	1.1736	2.1213	0.0338
typ/amp	L1	pass	female	1.283	0.782	1.6412	0.1007
typ/wc	L1	pass	female	0.711	0.5458	1.304	0.1921
typ/for	L2	pass	female	2.019	0.8904	2.2676	0.0233
typ/amp	L2	pass	female	-1.37	0.6937	-1.983	0.0473
typ/wc	L2	pass	female	-0.65	0.5575	-1.179	0.2380
typ/for	ASD	pass	female	0.278	0.3882	0.7172	0.4732
typ/amp	ASD	pass	female	0.200	0.4609	0.4348	0.6637
typ/wc	ASD	pass	female	0.024	0.409	0.061	0.9513
typ/for	L1	fail	male	1.508	1.0549	1.4301	0.1526
typ/amp	L1	fail	male	-0.53	0.5705	-0.934	0.3499
typ/wc	L1	fail	male	-1.61	0.973	-1.658	0.0971
typ/for	L2	fail	male	-0.64	0.5679	-1.132	0.2573
typ/amp	L2	fail	male	-0.00	0.7568	-0.010	0.9917
typ/wc	L2	fail	male	0.761	0.604	1.260	0.2076
typ/for	ASD	fail	male	0.110	0.796	0.1382	0.8900
typ/amp	ASD	fail	male	0.055	0.582	0.0949	0.9243
typ/wc	ASD	fail	male	-0.48	0.692	-0.698	0.4846
typ/for	L1	pass	male	-0.17	0.4954	-0.343	0.7314

typ/amp	L1	pass	male	-0.96	0.5047	-1.913	0.0556
typ/wc	L1	pass	male	-0.05	0.4013	-0.147	0.8829
typ/for	L2	pass	male	-0.75	0.4815	-1.575	0.1152
typ/amp	L2	pass	male	0.179	0.4623	0.3887	0.6974
typ/wc	L2	pass	male	0.793	0.5860	1.354	0.1757
typ/for	ASD	pass	male	2.460	1.2539	1.9623	0.0497
typ/amp	ASD	pass	male	-0.39	0.5724	-0.683	0.4944
typ/wc	ASD	pass	male	-1.57	0.704	-2.230	0.0257

As we would expect, these same age-related patterns were observed when running our analyses only on the participants who passed the reliability check (see Table 82).

Table 82. Post hoc trend analysis assessing age related change in likelihood of typical selection in relation to the visual comparison group, speech category and speaker gender for participants who passed the reliability check.

Visual Comparison	Speech Category	Speaker Gender	Age trend	*SE*	*z*	*p*
typ/for	L1	female	2.437	1.1242	2.168	0.030
typ/amp	L1	female	1.270	0.7664	1.657	0.0974
typ/wc	L1	female	0.625	0.5242	1.193	0.232
typ/for	L2	female	1.800	0.8477	2.123	0.033
typ/amp	L2	female	-1.283	0.6639	-1.93	0.053
typ/wc	L2	female	-0.606	0.5253	-1.15	0.2482
typ/for	ASD	female	0.293	0.3626	0.80	0.4181

typ/amp	ASD	female	0.119	0.4361	0.274	0.7834
typ/wc	ASD	female	0.052	0.381	0.138	0.889
typ/for	L1	male	-0.196	0.4664	-0.42	0.6742
typ/amp	L1	male	-0.945	0.4934	-1.91	0.0553
typ/wc	L1	male	-0.026	0.3779	-0.06	0.944
typ/for	L2	male	-0.678	0.466	-1.45	0.1458
typ/amp	L2	male	0.121	0.4371	0.278	0.7802
typ/wc	L2	male	0.731	0.5687	1.286	0.1983
typ/for	ASD	male	2.254	1.1949	1.886	0.059
typ/amp	ASD	male	-0.291	0.5392	-0.54	0.5889
typ/wc	ASD	male	-1.492	0.6857	-2.17	0.0294

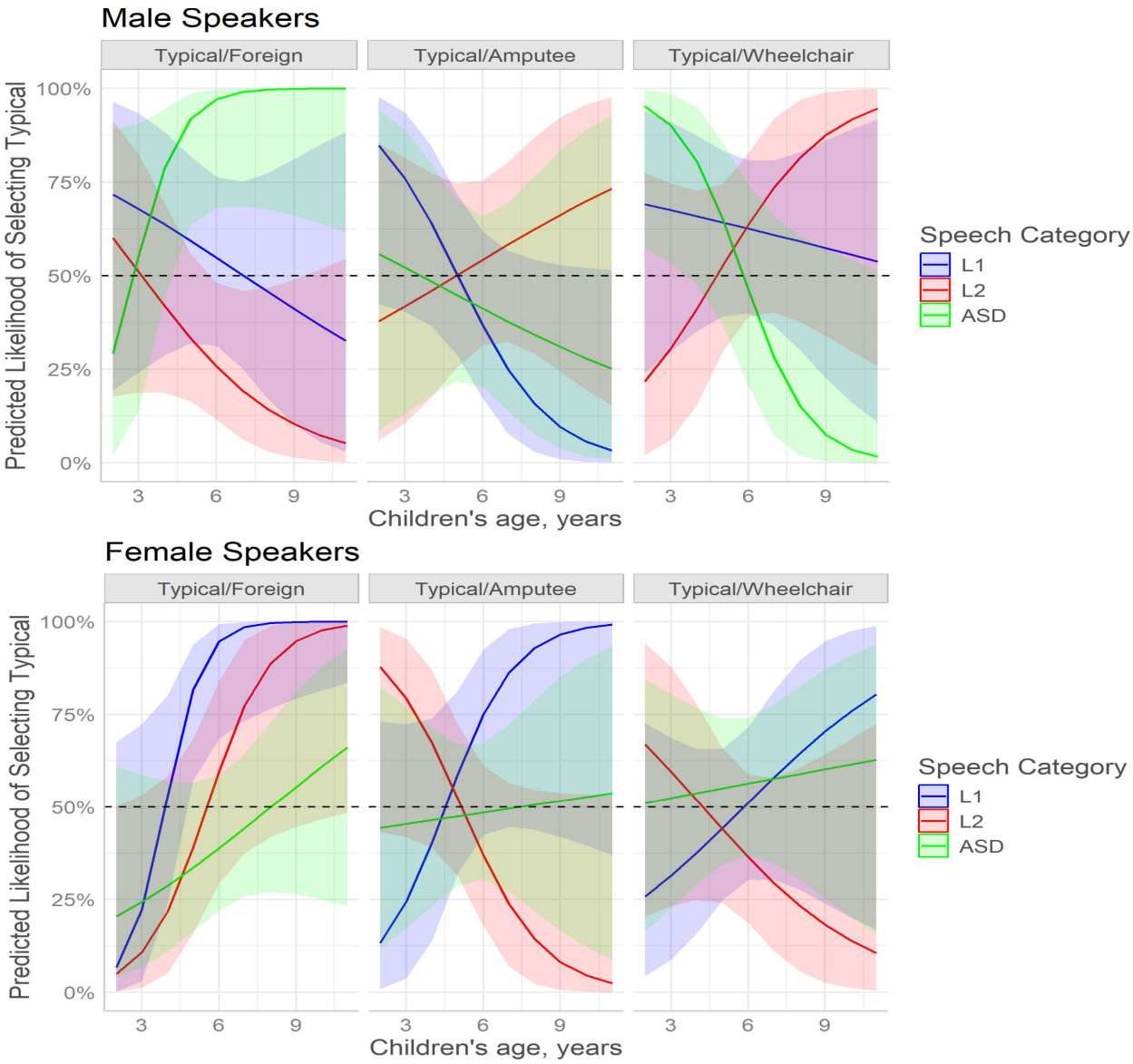

Figure 21. **Relationship between participants' age (in years) and predicted likelihood of selecting "Typical" images for male and female speakers. "Typical" refers to North American presenting images without disability, "Amputee" refers to North American presenting images with amputation, "WC" refers to North American presenting images of children in wheelchairs with contracture in the wrist and torticollis in the neck, and "Foreign" refers to foreign-presenting images without disability. Lines represent the predicted likelihood of selecting "Typical" for each visual comparison in relation to participants' age. Shaded areas around the lines represent confidence intervals based on the standard errors from the logistic regression.**

In sum, whereas children showed no measurable disability spread in the aggregate sample, an age-related increase in such tendencies when accounting for speaker gender as well as

participants' reliability check scores did emerge. Namely, with age, participants became increasingly likely to map male (but not female) speakers with ASD onto wheelchair (but not amputee) images. This suggests that, with age, participants became increasingly more attuned to features characterizing the (male) ASD speakers as well as the different features depicted in the amputee and wheelchair illustration categories, and that the wheelchair images were particularly likely to evoke disability spread tendencies.

The fact that, for children, disability spread was only apparent for male speakers is noteworthy. While future research is needed, we propose that underlying differences between the male and female ASD speakers used in this sample may account for this difference. Namely, in Redford et al.'s (2018) study, college-aged participants rated the speech of male ASD speakers used in the present study as significantly more disordered than the speech of the female ASD speakers used in the present study. Consistent with this finding, to recall, in study 1b, male ASD speakers (but not females) tended to receive higher dependence scores in comparison to L1 speakers. Thus, underlying differences in the severity of the speech disorder, rather than the gender of the speakers, per se, is likely the driver of the observed differences in disability-spread tendencies between male and female speakers with ASD in the present study. Notably, both in study 1b and in the present study, children who passed the reliability check were the participants who were most inclined to engage in disability spread vis-à-vis male ASD speakers (i.e., associating male ASD speakers with high dependence in study 1b and associating male ASD speakers with wheelchair images, with increasing age, in study 2b). While future research is needed, this suggests that the ability to engage in metacognitive reflection (such as those assessed by the reliability check) may be related to the tendency to relate ASD speech to a latent disability. Importantly, the large standard errors around the coefficients pertaining to age-related

change, as well as several anomalous findings in this regard, point to the need to replicate these findings with far larger sample sizes.

8.4.Summary

At the aggregate level, children did not associate disordered speech with physical disabilities in the way they associated foreign-accented speech with foreign garb. However, our exploratory analysis indicated significant age-related change in this regard. Namely, particularly for those who passed the reliability check, children displayed an age-related increase in the tendency to selectively associate male ASD speakers with wheelchair images, suggesting some degree of emerging analogous inductive generalization processes operating in children's evaluations of foreign and disordered speech.

We now turn to a discussion of similarities and differences in the associations children and college students drew between speech and appearance.

CHAPTER IX

STUDY 2 GENERAL DISCUSSION

In study 2, we assessed the extent to which participants would selectively associate speech produced by children with ASD with images depicting physical disabilities in a similar way to which they might associate foreign-accented speech with images depicting individuals dressed in foreign garb. This approach facilitated more direct exploration of participants' tendency to "spread" a perceived impairment in one domain (i.e., speech) to another (i.e., physical disability). As with study 1, the fact that children and college students were evaluated in the same way, using identical stimuli and prompts supported the ability to compare response patterns observed in these two samples in a straightforward manner. This allowed us to explore commonalities as well as meaningful differences between these two demographics.

Foreignness and Disability Related Inductive Generalizations

First, at the aggregate level, consistent with previous research (Wagner, Clopper and Pate, 2014), and with our pre-registered predictions, college students and children alike were more likely to associate L2 speech with foreign garb in comparison to L1 and ASD speech. In contrast, at the aggregate level, there was a clear difference between the college-aged and child samples vis-à-vis disability-related inductive generalizations. Namely, whereas college students clearly selectively associated ASD speech with wheelchair images, children did not. Thus, college students may have been more attuned to the features of ASD speech than children, thus eliciting greater globalized disability percepts. In support of this notion, there was a significant age-related increase in children's tendency to associate (male) ASD speakers with wheelchair images, suggesting that the tendency to associate ASD speech with a latent, globalized disability increases with age.

Along those lines, for both college students and children, the tendency to associate ASD speech with a disability was specific to wheelchair images but not amputee images. This pattern suggests that perceptions of the severity and kind of perceived impairments are highly relevant to the expression of disability spreading tendencies. Concretely, impairments perceived as non-localized may elicit greater inductive generalizations in comparison to impairments perceived as localized.

Notably, whereas ASD/wheelchair associations were only evident for older children evaluating male ASD speakers, college students associated ASD speakers with wheelchair images irrespective of the gender of the speakers. This suggests that college students may have been more attuned to even the more subtle indicators of a speech disorder in the female ASD speakers.

Summary

In sum, the results from study 2 shed further light on inductive generalizations, relations between foreign and disability related generalizations, and age-related change in this regard. Whereas the college students clearly mapped ASD speech onto wheelchair images in a similar manner to which they mapped L2 onto foreign garb (and perhaps even to a greater extent), the results were more mixed for children. Nevertheless, there was an emerging age-related increase in this tendency for children as well.

CHAPTER X

GENERAL DISCUSSION

10.1. Overview

In this dissertation, we took a comparative approach to assessing the development of foreignness and disability concepts. Our overarching goals were to shed light on 1) children and college students' tendency to discriminate between foreignness and disability cues, 2) differential generalizations stemming from these discriminations, and 3) expression of prejudice towards foreign individuals and individuals with a disability. Our novel methodological approach afforded multiple comparative analyses, including across social concepts (disability and foreignness), modalities (speech and appearance), and ages (wide age span of children as well as emerging adults). This allowed us to glean highly articulated information about perceptual cues in speech and appearance that children and college students are attuned to, and the socially significant inferences they draw based on these cues.

In what follows we provide a brief re-cap of our central findings and discuss how these relate to the existing literature. We will then discuss a series of methodological limitations related to the stimuli used in this dissertation and the way we operationalized key variables. Finally, we end with a discussion of ways in which our findings can be leveraged towards potential avenues for intervention aimed at reducing the expression of prejudice towards stigmatized others.

10.2. Summary of Key Findings

10.2.1. Conceptual Differentiation

In study 1, we found that children and college students alike selectively associated images depicting physical disabilities with low physical competence and high dependence. They

also selectively associated foreign garb and foreign accent (L2) with being "from another country far away" but not with low competence or high dependence. As well, in study 2, children and college students were both inclined to associate L2 speech with foreign garb. Such findings reflect ways in which children and college students exhibited common inferential assumptions regarding the perceptual variability they evaluated.

The findings from this dissertation also replicate and extend previous work assessing children's sociolinguistic awareness. Specifically, children's tendency to explicitly associate L2 speech with foreignness was consistent with findings reported by Kinzler and Dejesus (2013); children's tendency to map L2 speech onto foreign garb in the present research was generally consistent with previous research as well (e.g., Hirschfeld and Gelman, 1997; Wagner, Copper and Pate, 2014; Weatherhead, White and Friedman, 2016). Such replications are encouraging in terms of the integrity and reliability of the present studies and highlight the robustness of these previous findings as well.

Notably, there were also important differences between college students and children's tendency to differentiate between perceptual cues to disability and foreignness. To recall, in study 1, children (but not college students) were somewhat unsure regarding the extent to which children depicted in wheelchairs with signs of contracture in the wrist and torticollis in the neck, as well as speakers with ASD, were "from America" or "another country far away". The ambiguous foreignness judgments children provided point to ways in which their differentiation between foreignness and disability concepts may yet be limited. On the other hand, in study 2, children and college students were both disinclined to associate L2 speech with images depicting physical disabilities and did not associate ASD speech with foreign garb. Thus, future research

is needed to assess the extent to which the ambiguity observed in children's responses in study 1 was due to the experimental design or to protracted conceptual differentiation.

10.2.2. Inductive Generalizations related to Foreignness and Disability

The ability to assess relations between inductive generalizations in the case of foreignness and disability was a key advantage of our comparative design. Notably, we observed important differences between children and college students in this regard. To recall, in study 2, college students associated ASD speakers with wheelchair images (but not amputee images) to a similar degree to which they associated L2 speakers with images depicting foreign garb (and perhaps even to a greater extent). In contrast, at the aggregate level, children were not inclined to associate ASD speakers with either of the images depicting physical disabilities. Thus, college students' responses point to inferential processes at play in relation to perceptions of foreignness and disability. In both cases, perceptual cues detected in one domain (i.e., speech: speech disorder; accent) were taken to license inductive inferences about another domain (i.e., appearance: physical disability; foreign garb). To our knowledge, this is the first study to directly uncover these cross-domain inductive processes. While further research is needed, one potential implication of these findings is that people may develop intuitive theories about disabilities in a similar manner to which they develop intuitive theories about other social groups such as race (Mandalaywala et al., 2019) and gender (Gelman and Taylor, 2014). Namely, a latent concept (i.e., "foreign"; "disabled") appears to have warranted a collection of predictions regarding people categorized in relation to these social concepts.

Previous researchers have implicated inductive generalization processes related to group perception in the expression of ethnic and racial prejudice, going so far as describing these

inductive generalizations as the "original sin of cognition" (Leslie, 2017:421). There is an urgent need for further research that continues to explore the nature of inductive generalizations in the case of disability and the relation between these generalizations and expression of prejudice. One implication of our data for this important area of inquiry is the possibility of developmental change in relationships between inductive inference and prejudice. To recall, in study 1, the tendency to associate stimuli containing disability cues in speech or appearance with enhanced dependence and reduced physical competence (indices of inductive inference that we operationalized as disability spread) was generally not associated with the tendency to provide lower friendship scores to children with ASD speech or with children in images with physical disabilities. In fact, for female children, perceptions of dependence were associated with enhanced interest in friendship with children in wheelchairs with signs of severe cerebral palsy. In stark contrast to the patterns observed with children, disability spread was generally associated with college students' expression of prejudice towards disability-related stimuli (see chapter 5 for discussion). Thus, we observed a developmental shift from a) in children, a neutral (and sometimes positive) relationship between inductive generalizations and prosocial attitudes towards individuals with a disability, to b) the pattern observed in the college sample (in which inductive generalizations were associated with prejudice).

Notably, in contrast to the aggregate child sample, children (particularly those who passed the reliability check) did display an age-related increase in the tendency to selectively associate male ASD speakers with wheelchair images. This suggests some degree of emerging inductive generalization processes operating in children's evaluations of foreign and disordered speech. Thus, the period may be fleeting in which children do not readily associate an observed impairment with a globalized latent disability in the way that adults do. This may have

implications for the timing of interventions aimed at reducing negative attitudes towards individuals with disabilities (as we will discuss below).

10.2.3. Stigma is Dynamic

Whereas several of our findings were consistent with the extant literature, as noted, we were particularly surprised to find that children in our study showed no preference for native-accented (L1) speech over foreign-accented (L2) speech. We found this to be surprising due to the wide range of generic statements made in the developmental literature regarding children's preference for native-accented speakers. For example, Kinzler and DeJesus (2013:655) note that: "By 5– 6 years of age, monolingual children express robust social preferences for native-accented speakers of their native language." Similarly, Paquette-Smith et al., (2019:809) note that "By five years of age, children already show strong social preferences for peers who speak their native language with a familiar accent" (see also Kinzler et al., 2007; Kinzler et al., 2009; Souza., Byers-Heinlein and Poulin-Dubois, 2013).

That said, previous accounts have explicitly acknowledged that native-accented preference is not absolute. For example, Kinzler and DeJesus (2013) found that children preferred foreign-accented speakers who described prosocial actions they have committed (e.g., "helping someone up the playground") over native-accented speakers who described antisocial actions they have committed (e.g., "pushing someone on the playground"). However, in such studies factors such as a speaker's perceived "niceness" or "meanness" are described as factors that may mitigate a more "primary" (Kinzler, Shutts and Correll, 2010:584) attunement to native-accented speech. As Kinzler et al., (2009:632) surmised: "Given that social groups in ancient times likely differed in accent, but not in race, children may be predisposed to rely primarily on accent to guide their social evaluations of novel individuals."

While descriptions of a primary preference for native-accented speech may seem compelling from evolutionary perspectives, they may fall short from capturing the complex and multi-faceted nature of speech processing. Namely, factors such as perceivers' openness to experience, social dominance orientation, attitudes toward the particular foreign accent spoken, and the more general sociolinguistic cultural context (i.e., harmony vs. inter-group tension) in which perceivers exist all likely exert significant influence on the way a foreign accent will be received. There is therefore a great need to incorporate this sort of complexity into theoretical accounts describing children's developing sociolinguistic proclivities.

Indeed, from a cultural psychology perceptive, a complete theoretical account of any given type of stigma must involve accounting for explanatory concepts used to interpret human diversity that vary significantly across time and place (Koschorke et al., 2017; Munyi, 2012). To illustrate, the so-called "hearing aid effect" (which we described in our introduction as an example of disability spread) is a term used to describe the tendency to draw negative attributions regarding individuals using hearing aids. While this effect was first empirically identified by Blood and colleagues in 1977, the extent to which individuals in the United States continue to exhibit this tendency in the 21st century is not clear (see Rauterkus and Palmer, 2014, though see also Wheeler and Tharpe, 2020). The dynamic nature of the hearing-aid effect illustrates the need to consider cultural context in theoretical and empirical investigation of stigma – both in the case of foreignness and disability.

This dissertation further highlights the need to consider developmental change as another aspect of dynamic change in stigma. For example, whereas prejudice against L2 speakers was absent at all ages in the present findings, college students (but not children) provided significantly lower friendship scores for ASD speakers in comparison to the other speech

categories, consistent with previous research (Redford et al., 2018). Conversely, children were prejudiced against the wheelchair (and foreign) images, while college students showed a prosocial bias towards images depicting physical disabilities (particularly amputees). Clearly, there is a need for future research to explore age-related change in the expression of prejudice and reasons why such expressions of prejudice may be expressed in different ways across the lifespan.

10.2.4. Advancing Theoretical Understanding of Disability Spread

As noted, both children and college students displayed a tendency to associate ASD speech with a disability, but this was specific to wheelchair images (and did not extend to amputee images). This pattern suggests that perceptions of the severity and kind of perceived impairments are highly relevant to the expression of disability spreading tendencies. Concretely, impairments perceived as non-localized (such as the wheelchair images in the present study) may elicit greater inductive generalization in comparison to impairments perceived as localized. The selectivity with which participants engaged in disability spread raises questions regarding our theoretical understanding of this phenomenon. Namely, in Wright's (1964) original formulation, disability spread was described as a negatively valanced halo-effect. That is, perceivers were said to engage in disability spread because it is "cognitively easier" (p. 198) to form a generalized negative perception of individuals with a disability rather than form more complex representations that include both positive and negative components. Along those lines, disability spread was said to be "difficult to escape" (p. 198) due to this relative cognitive ease, consistent with long-standing approaches in psychology that have described humans as "cognitive misers" (see Stanovich, 2020 for discussion).

A potential challenge to such accounts is that, rather than reflecting non-rational, miserly cognitive processes, disability spread may instead be firmly rooted in reasonably accurate representations of statistical distributions in the real world. To illustrate, motor deficits are common in children with ASD, even within the first year of life (West, 2019). As well, population-based studies have found that speech difficulties affected more than one in two children with cerebral palsy (Nordberg et al., 2013), while other studies have found the rate of co-occurrence to be as high as 82% (Mei et al., 2020). Considering these data, college-aged participants' tendency to associate ASD speakers with reduced physical competence (study 1) or with a physical disability evoking severe cerebral palsy (study 2) can be seen as potentially reflecting statistically valid inferences regarding the co-occurrence of impairments in disabilities such as ASD and cerebral palsy. To return to Wright (1964), disability spread may therefore be "difficult to escape" due to the statistical validity of such inferences rather than factors such as cognitive ease.

While future research directly examining this issue is sorely needed, we suspect that disability spread typically emerges in relation to essentialist interpretive frameworks that extend beyond mere statistical learning. To illustrate, whereas essentialist and non-essentialist perceivers alike may engage in disability spread, the interpretive framework giving rise to these inductive processes may differ significantly. Namely, whereas essentialists will likely infer that a latent disability essence caused a range of observed impairments, non-essentialists will be more likely to grasp that co-occurring impairments may emerge due to a range of dynamic processes rather than a singular disability essence. Thus, the essentialist interpretive framework is likely to drive the perception that an individual with a disability is fundamentally disabled (Dunn, 2015).

For example, motor deficits in ASD could impact affected children's ability to actively engage with their social environment. In turn, this can result in reduced social interaction and opportunities for language growth (see Reindal et al., 2022). Speech impairments observed in the case of ASD may therefore be traced to cascading developmental processes rather than to a singular essential disability per se. Whereas such accounts of the relationship between speech impairment and ASD may be readily endorsed by perceivers operating with a non-essentialist interpretive lens, essentialist perceivers may be more likely to simply attribute any observed impairment to a singular disability essence, as noted (see Chapman, 2020).

Since non-essentialist accounts place the causal force driving the expression of disabilities on the interaction between the organism and their environment rather than the essential features of the individual, they may result in reduced stigma towards those individuals and greater interest in improving outcomes. The potential prosocial consequences of non-essentialist responses to individuals with co-morbid disabilities underscore the importance of future research assessing the extent to which essentialist interpretive frameworks dominate children's (and adults') disability concepts in ways that resemble essentialism regarding concepts such as race or gender.

Conversely, future research should also explore the possibility that essentialist interpretive frameworks may also mitigate the expression of prejudice towards individuals with co-morbid disabilities. For example, Bogart, Rosa and Slepian, (2019) found that essentialism may ameliorate prejudice towards individuals with acquired disability or a chronic illness like HIV by refuting the sense that these conditions are controllable and blameworthy (see also Peretz-Lange, 2021). Such findings underscore the need for future research to directly examine

the relationship between essentialist interpretive frameworks regarding co-morbid disabilities and the expression of prejudice.

10.3. Limitations

While we have highlighted the many strengths of our work, it is also important to acknowledge several limitations in our general approach, our stimuli and the way we operationalized key variables.

10.3.1. Sample size

First, while our sample size was larger than typical studies assessing children's sociolinguistic awareness, we were still under-powered to reliably assess two- and three-way interactions. Thus, the results of our moderation analysis (including age-related change) should be interpreted as only reflective of tendencies that were apparent in our sample. Future research with far greater sample sizes is needed to assess the extent to which these findings are replicable and generalizable.

10.3.2. Visual stimuli limitations

Whereas the speech stimuli used were naturalistic with optimal external validity (i.e., L2 speech was in fact produced by children living in Spain – a country "far away"), the illustrations were caricatured depictions which relied on stereotypical notions (i.e., children in Spain generally do not, in fact, dress like matadors). Thus, there is a need for future research to assess whether our findings hold even for real-life visual depictions of foreignness and disability. Notably, the fact that our stereotypical stimuli evoked robust foreignness judgements in the current study is consistent with prior work indicating that stereotypical (and often highly inaccurate) intuitions are common in folk sociology (e.g., associating Portuguese speech with an igloo; Hirschfeld and Gelman, 1997). Our findings thus suggest that, at least under certain

circumstances, stereotypical cues (even inaccurate ones) can be as potent as realistic ones in driving social categorization processes, though future research is needed (see Bian and Cimpian, 2017).

10.3.3. Speech stimuli limitations

While the speech stimuli used in our studies were far closer to realistic depictions than our visual stimuli, some limitations were nonetheless present. First, the irregular features characterizing the ASD speech samples used in this study were relatively mild. On the one hand, this fact increases the noteworthiness of children's ability to discriminate ASD speech from both L1 and L2 speech, as well as of the signs of disability spread we observed in both children and college students. That said, the extent to which the patterns we observed would generalize to instances of more severe speech disorders is not clear. Future research exploring participants' reactions to speech disorders ranging in kind and severity is therefore needed. Second, when listening to the unedited recording sessions in which children with ASD engaged in storytelling to an adult, it appeared that several of these children were not particularly happy about the activity. In fact, some children with ASD appeared to need encouragement to complete the story-telling activity (though participants in our study only heard the uninterrupted speech produced by autistic children, without those words of encouragement). It is therefore possible that frustration "leakage" with the storytelling task may have been expressed in autistic children's speech patterns. For example, they may have sounded tired or exasperated due to being prodded to complete the story. This raises questions regarding the most appropriate way to interpret participants' evaluations of these speakers. That is, it is possible that these speakers were associated with reduced physical competence or enhanced dependence due to situational factors during stimuli collection rather than stable features of their speech. On the other hand,

ASD speakers used in the study were all described by speech language pathologists as having irregular prosody, indicating that their speech had atypical features that extended beyond the context in which they were recorded (see Redford et al., 2018). Regardless, future work using even more naturalistic speech could clarify these issues. A third limitation relating to the speech stimuli used in the present study is that L2 speech was obtained by different researchers with different (though similar) storytelling details and elicitation procedures than the L1 and ASD speech. Thus, our study did not meet the strict standards typically employed in matched-guise experiments in which the same speaker recites the same words in different trials (i.e., French vs. English accent) (Kircher, 2015; Nejjari et al., 2019). On the other hand, matched-guise techniques have their own limitations as well (Garrett, 2010), and experiments assessing evaluations of speech varieties often compare stimuli elicited in somewhat different contexts (e.g., Fernald, 1985; Hilton, 2021; see Weinstein and Baldwin, 2024 for discussion). Thus, while our approach was justified as a first-of-its kind comparison between ASD and L2 speech, future research should gather speech stimuli using identical prompts.

10.3.4. Operationalization Limitations

Beyond the limitations of our visual and speech stimuli, there were several limitations associated with the way we operationalized our variables in study 1. First, our foreignness prompt relied on explicit knowledge of the idea of there being countries different from "America" that are "far far away," which young children may have struggled with. That said, asking whether someone is "from America or another country far away" in a binary manner suggests that being "from America" contrasts with being from "far away". Thus, to the extent that a given stimulus evoked a sense of being from "far away", young children's ability to

engage in pragmatic inference should have supported their ability to comprehend this prompt (see Horowitz and Frank, 2016; Baharloo et al., 2023 who make a similar case).

Second, our assessment of peer rejection involved asking participants whether they were interested in being friends with targets – an approach that was limited for two main reasons. First, due to social desirability bias (which children are also prone to – see Miller et al., 2015), participants may have been reluctant to explicitly reject interest in befriending someone. Second, with this method, peer rejection could only be inferred by low interest in friendship. Thus, future work should consider ways to overcome potential desirability bias while also measuring prejudice more directly. However, since there was variability in children and college students' responses to the friendship prompt, and some degree of peer rejection was in fact detected, our approach seems to have been justified.

In addition to potential issues with the way we operationalized peer rejection towards individuals with a disability, our discussion of predictors of such peer rejection was limited as well. Namely, in contrast to our framework (chapter 5) which was rooted in attribution theory (Weiner, 2005;2006) previous accounts of disability prejudice have described aversive reactions to individuals with a disability as stemming from an over-active disease avoidance mechanism (Nario-Redmond, 2024). According to such accounts, non-disabled individuals may avoid interacting with individuals with a disability due to an intuitive fear of contagion. Since evolutionary approaches to disability prejudice describe these contagion-avoidance reactions as unlearned, there is a great need to empirically evaluate whether children exhibit contagion fear when encountering individuals with a disability. In sum, future work on this topic should adopt a broader approach to assessing the emergence of disability-related prejudice and peer rejection in childhood.

10.3.5. Ethical Limitations

Lastly, is also important to consider a potential ethical drawback to the design employed in study 2. To recall, in study 2, we asked participants to map speech varieties onto appearance. While the prompt was framed as merely probing the participants' opinion regarding which image likely produced the speech, it is entirely possible that children may have also inferred that there was a correct answer (see Westra and Carruthers, 2017). If so, the experimental design may have inadvertently communicated to children that they ought to associate children who look a certain way (e.g., with physical disabilities) with a certain type of speech (see also Gardner and Lambert, 1972). It is therefore important that future assessment of children's disability concepts consider ways in which experimental designs may inadvertently communicate potentially harmful messages, and if so, find ways to mitigate this problem.

10.4. Implications for interventions

10.4.1. Window of Opportunity - Perceptual attunement may precede prejudice

To recall, although children did not associate ASD speakers with physical disabilities in study 2 the way college students did (at the aggregate level), children did discriminate between ASD speech and the other speech categories vis-à-vis foreignness and comprehensibility judgments in study 1 (though this pattern was more apparent for children who passed the reliability check). Furthermore, in contrast to college students, children in study 1 did not exhibit prejudice against ASD speakers in their friendship ratings (irrespective of their reliability check scores). This suggests that children's perceptual attunement to ASD speech preceded their negative evaluative tendencies regarding the speakers producing such speech. Notably, the fact that the reliability-check scores predicted perceptual attunement to disability features in speech

vis-à-vis the foreignness, comprehensibility and competence ratings, but was not significantly related to friendship scores, supports the claim that perceptual attunement may precede prejudice as well.

The disassociation we observed between children's perceptual attunement and evaluative judgments vis-à-vis speakers with ASD points to a potential window of opportunity for interventions aimed at reducing prejudice towards individuals with a speech disorder. That is, prior to children exhibiting perceptual attunement to the features of disordered speech, interventions may be futile since children may not even detect the variability at hand. As well, once prejudicial attitudes are formed, they may be relatively resistant to change (Killen et al., 2011; Lee et al., 2017). Thus, interventions targeting children who are already perceptually attuned to the features of disordered speech but have not yet come to develop negative attitudes towards speakers with a speech disorder may be particularly promising. However, caution is warranted before developing interventions that target such a window of opportunity, as we discuss below.

10.4.2. Should Disability Spread be Combatted or Harnessed?

The effectiveness of any intervention to reduce disability-related stigma will depend, in part, on adequate understanding of the relationship between disability spread and prejudice. As noted, disability spread was associated with the expression of prejudice for college students, but not for children. Furthermore, for female children, perceptions of dependence were associated with enhanced interest in friendship with children in wheelchairs with signs of cerebral palsy. These results raise the possibility that interventions geared towards modifying children's tendency to attribute reduced competence to individuals with disabilities may inadvertently compromise children's prosocial impulse to provide care to those perceived of as being in need.

That is, to the extent that perceptions of dependence increase prosocial motivations, challenging the notion that individuals with a disability are dependent may ultimately result in reduced prosocial motivations towards individuals with a disability as well.

Future research should therefore assess the extent to which interventions that exclusively highlight the relative strengths of individuals with a disability (thus presumably decreasing spreading tendencies) are more or less effective at a) improving prosocial attitudes in comparison to interventions which highlight those strengths, while b) also nurturing children's impulse to provide care, compassion and comradery to those they perceive as being in need.

10.5. Conclusion

Children as well as adults readily draw generalizing inferences about others based on minimal exposure. While such inductive inferences reflect basic cognitive processes that form the backbone of all conceptual representations, they may also be associated with the expression of prejudice towards stigmatized others. In this study, we assessed children and college students' tendency to differentiate between perceptual cues to foreignness and disability cues, the inductive generalizations drawn based on this conceptual differentiation, and the way these inferences relate to the expression of prejudice. While college students and children expressed similar levels of conceptual differentiation (with some exceptions), they differed significantly in the inductive generalizations drawn based on this conceptual differentiation and in their expression of prejudice. This work advances our understanding of children's developing conceptual understanding of foreignness and disability alike. Our findings also hold potential to inform development of empirically oriented interventions to reduce the expression of prejudice in childhood and across the lifespan.

Supplementary Materials

1. Visual Stimuli.

2. Demographics

Study 1a:

Sociodemographic Characteristics of Participants

Sample Characteristics	*n*	*%*	*M*	*SD*
Gender				
Male	50	27.8		
Female	126	70.0		
Gender Fluid	1	0.5		
Transgender	2	1		
Other	1	0.5		
Race				
Asian	16	8.8		
White	138	76		
Alaska Native or IA**	2	1.1		
Black or African American	11	6.1		
Other	11	6.1		
Prefer not to say	2	1.1		
Languages Spoken at Home				
1	113	62.7		
2	59	32.7		
3	8	4.4		
Disability Exposure				
Exposure to a Child	43	23.9		
Exposure to an Adult	91	50.5		
Ethnicity				
Not Hispanic or Latino	150	83.3		
Hispanic or Latino	30	16.7		
Age (in years)			19.73	2.04
Parental Education***			15.74	3.07

* Pacific Islander

** Indigenous American

*** In years of education

Study 1b:

Sociodemographic Characteristics of Participants

Sample Characteristics	*n*	*%*	*M*	*SD*
Gender				
Male	73	44.8		
Female	89	54.6		
Other	1	0.6		
Race				
Asian	18	11		
White	108	66		
Native Hawaiian or PI*	3	1.8		
Alaska Native or IA**	4	2.5		
Black or African American	11	6.7		
Other	18	11		
Prefer not to say	1	0.006		
Languages Spoken at Home				
1	100	61.3		
2	59	36.2		
3	4	2.5		
Disability Exposure				
Exposure to a Child	23	14.1		
Exposure to an Adult	26	16.0		
Ethnicity				
Not Hispanic or Latino	141	86.5		
Hispanic or Latino	22	13.5		
Age (in years)			5.7506	1.6075
Parental Education***			16.9	3.2905

* Pacific Islander

** Indigenous American

*** In years of education

Study 2a:

Sociodemographic Characteristics of Participants

Sample Characteristics	*n*	*%*	*M*	*SD*
Gender				
Male	47	34.5		
Female	84	61.7		
Gender Fluid	3	2.2		
Transgender	2	1.4		
Race				
Asian	17	12.5		
White	101	74.2		
Native Hawaiian or PI*	1	0.007		
Alaska Native or IA**	4	2.9		
Black or African American	7	5.1		
Other	6	4.4		
Languages Spoken at Home				
1	84	61.7		
2	46	33.8		
3	6	4.4		
Disability Exposure				
Exposure to a Child	43	31.6		
Exposure to an Adult	78	57.35		
Ethnicity				
Not Hispanic or Latino	117	86.02		
Hispanic or Latino	19	13.9		
Age (in years)			19.75	1.002
Parental Education***			16.4	3.08

* Pacific Islander

** Indigenous American

*** In years of education

Study 2b:

Sociodemographic Characteristics of Participants

Sample Characteristics	*n*	*%*	*M*	*SD*
Gender				
Male	68	47.5		
Female	75	52.4		
Race				
Asian	21	14.68		
White	87	60.8		
Native Hawaiian or PI*	3	2.09		
Alaska Native or IA**	2	1.4		
Black or African American	6	4.1		
Other	18	12.58		
Prefer not to say	6	4.1		
Languages Spoken at Home				
1	86	60.1		
2	53	37.06		
3	4	2.8		
Disability Exposure				
Exposure to a Child	19	13.4		
Exposure to an Adult	21	14.9		
Ethnicity				
Not Hispanic or Latino	121	84.6		
Hispanic or Latino	22	15.4		
Age (in years)			5.408	1.8875
Parental Education***			16.62	3.43

* Pacific Islander

** Indigenous American

*** In years of education

3. Open-Ended Analysis

Our primary interest in the open-ended responses was whether children would spontaneously refer to disability or foreignness in their initial encounter with the stimuli. More generally, we were also interested in children's spontaneous evaluations of the stimuli (e.g., causal inferences, affective expressions, etc.) and the factors they found most salient in each stimuli category. Overall, we obtained intelligible responses from 70 children (for the audio stimuli) and 126 children (for the visual stimuli).

Speech

Of the 54 comments received for L2, 8 (14.8%) referred to the content of the speech, 13 (24%) made explicit reference to nationality or language/accent (e.g., "It's in Spanish", "I think it's Korea"), 7 (13%) referred to comprehensibility (e.g., "can't understand them", "it's easy to understand them) and 4 (7.4%) made evaluative comments (e.g., "silly voice", "sounded weird").

Of the 52 comments received for ASD, 20 (38.4%) referred to the content of speech, 3 (5.8%) mentioned nationality or language/accent (e.g., "sounded like he's from Mexico", "It's Korean"), 4 (7.7%) referred to comprehensibility (e.g., "can't understand", and 7 (13.5%) made evaluative comments ("stumbled over words", "they speak differently").

Of the 64 comments received for L1, 31 (48.4%) referred to the content of the speech, 5 (7.8%) made reference to nationality or language/accent (e.g., "from America", "sounds like she is American"), 4 (6.2%) made reference to comprehensibility ("understood well", "I can hear them well"), and 7 (10.9%) made evaluative comments (e.g., "talks normal", "good reader").

Visual

Of the 108 comments received for WC images 37 (34.2%) referred to a wheelchair (e.g., "can't walk needs wheelchair") and 21 (19.4%) referred to sitting (e.g., "watching a movie", "relaxing in a chair", "I think they're in a plane"). 5 (4.6%) referred to the hand ("he looks like he broke his hand", "his hand is like this"), 9 (8.3%) referred to mental states (e.g., "bored", "happyish", "sadish"), and 10 (9.26%) referred to a disability without mentioning a wheelchair (e.g., "she looks hurt", "their muscles are not moving well", "she has a big disability"). The comments thus reflect the relative salience of the wheelchair (more so than the neck or hand). The large proportion of references to sitting without acknowledging the wheelchair reflects the spectrum of understanding of disability in our sample.

Of the 111 comments received for amputee images, 39 (35%) mentioned a missing leg/foot ("doesn't have a leg"), and 28 (25%) referred to crutches/sticks though some did not seem to understand the significance of the crutches (e.g., "running and has sticks"). 19 (17%) referred to an injury ("broken leg", "his knee is broken"). While some assumed an accident (e.g., "his leg got cut off"), others assumed innateness (e.g., "he was born without a leg"). As with the wheelchair stimuli, responses reflect the salience of the disability cue to children while also reflecting the diverse interpretations of the disability (e.g., injury, permanence) and range of conceptual understanding (e.g., whether the "sticks" relate to a disability).

Of the 94 comments received for the foreign images, 38 (40%) referred to the outfit (e.g., "wearing dress", "dressed up like a pirate"), 9 referred to fanciness (9.5%) (e.g., fancy), 3 made evaluative comments ("looks like a showoff", "looks handsome", "cool"). 2 referred to foreign nationality ("he looks like he's from China", "I think the dress is from China"), and 2

referred to difference (“dressed different than normal”, “she looks different”). Thus the dress was the most salient feature to children but most attributed it to factors other than nationality (e.g., pirate, fancy).

Of the 54 comments received for typical images, 9 (16.6%) referred to normality (e.g., “he’s normal”, “he’s fine”), 8 (14.8%) referred to the posture (e.g., “standing”). And 10 (18%) referred to factors such as gender/clothing/hair. Thus, no factor dominated children’s comments in contrast to the other stimuli categories.

4. Gender Differences in Stimuli

4.1.Foreignness Judgment (Visual)

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
Family: binomial (logit)
Formula:
prob_select_faraway ~ Image_type * image_gender + (1 | participant_private_id)
Data: location_visual

AIC BIC logLik deviance df.resid
880.1 920.5 -431.1 862.1 644

Scaled residuals:
Min 1Q Median 3Q Max
-2.2525 -0.8427 0.4440 0.8869 1.4312

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.427	0.6534

Number of obs: 653, groups: participant_private_id, 167

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.1759	0.2368	-0.742	0.457
Amputee	-0.2311	0.3214	-0.719	0.472
Wheelchair	0.3416	0.3233	1.0566	0.290
Foreign	0.6991	0.3260	2.1444	0.031
Male	-0.142	0.3467	-0.409	0.681
Amputee x Male	0.4632	0.4687	0.9881	0.323
Wheelchair x Male	0.0055	0.4694	0.0117	0.990
Foreign x Male	0.9976	0.4965	2.0092	0.044

simple contrasts for Image_type

contrast	image_gender	estimate	SE	z.ratio	p.value
Typical - Amputee	Female	0.23112	0.32140	0.7191	0.88953
Typical - WC	Female	-0.3416	0.323303	-1.056	0.71591
Typical - Foreign	Female	-0.6991	0.326024	-2.144	0.13917
Amputee - WC	Female	-0.5727	0.325189	-1.7612	0.29219
Amputee - Foreign	Female	-0.9302	0.32833	-2.833	0.02383
WC - Foreign	Female	-0.3575	0.3275	-1.0916	0.69466
Typical - Amputee	Male	-0.2321	0.34095	-0.6808	0.90442
Typical - WC	Male	-0.3471	0.341041	-1.0178	0.73896
Typical - Foreign	Male	-1.6968	0.379372	-4.4726	0.00004
Amputee - WC	Male	-0.1150	0.33925	-0.3389	0.986594

Amputee - Foreign	Male	-1.464673	0.376296	-3.8923	0.00057
WC - Foreign	Male	-1.34966	0.37542	-3.59508	0.00184

4.2. Dependence Judgments (Speech)

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Dependence ~ audio_type * speaker_gender + (1 | participant_private_id)
Data: help
REML criterion at convergence: 510.3

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.9677	-0.6637	0.1845	0.6257	1.9723

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.2843	0.5332
Residual		0.4319	0.6572

Number of obs: 213, groups: participant_private_id, 71

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	1.925	0.1338	157.4	14.3861	<0.00001
L2	0.1749	0.1469	138.0	1.190803	0.23577
ASD	0.0499	0.1469	138.0	0.34022	0.7342
Male	-0.054	0.2025	157.4	-0.26681	0.7899
L2 x Male	0.0830	0.2224	138.0	0.37348	0.7093
ASD x Male	0.33709	0.2224	138.0	1.5156	0.1318

`simple contrasts for audio_type`

contrast	speaker_gender	estimate	SE	df	t.ratio	p.value
L1 - L2	female	-0.1749	0.1469	137.9	-1.1908	0.460
L1 - ASD	female	-0.0499	0.1469	137.9	-0.340	0.9382
L2 - ASD	female	0.125	0.1469	138	0.8505	0.6722
L1 - L2	male	-0.258	0.1669	137.9	-1.545	0.272
L1 - ASD	male	-0.387	0.1669	137.9	-2.318	0.056
L2 - ASD	male	-0.129	0.1669	138	-0.772	0.72

5. Study 1a Aggregate Sample Analysis

5.1. Conceptual differentiation:

To what extent will college students selectively associate physical disability cues (i.e., amputation, wheelchair) but not foreignness cues (i.e., foreign garb) with reduced physical competence (running ability)?

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Competence ~ Image_type + (1 | participant_private_id)
Data: running

REML criterion at convergence: 806.6

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.7859	-0.4485	-0.2035	0.0260	4.3524

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.02494	0.1579
Residual		0.16655	0.4081

Number of obs: 680, groups: participant_private_id, 170

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.28235	0.033	643.2	68.004	<0.00001
Amputee	-1.14117	0.044	506.9	-25.78	<0.00001
Wheelchair	-1.19411	0.044	506.9	-26.9763	<0.00001
Foreign	-0.24117	0.0442	506.9	-5.44842	<0.00001

To what extent will college students selectively associate foreignness cues (i.e., foreign accent and garb) but not disability cues (i.e., amputation, wheelchair) with being from another country?

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)
[glmerMod
]
Family: binomial (logit)
Formula: prob_select_faraway ~ Image_type + (1 | participant_private_id)
Data: location_visual

AIC	BIC	logLik	deviance	df.resid
492.9	515.5	-241.4	482.9	675

Scaled residuals:

Min	1Q	Median	3Q	Max
-18.011	-0.033	-0.013	0.002	78.587

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	97.66	9.882

Number of obs: 680, groups: participant_private_id, 170

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-6.8839	0.755	-9.1177	<0.00001
Amputee	0.0000008	0.542	0.000001	0.9999
Wheelchair	-1.914486	0.677	-2.82749	0.0047
Foreign	12.595949	1.2431	10.13262	<0.00001

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod
]
Family: binomial (logit)
Formula: prob_select_faraway ~ audio_type + (1 | participant_private_id)
Data: location_audio

AIC	BIC	logLik	deviance	df.resid
366.3	383.3	-179.2	358.3	506

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.1299	-0.3040	-0.1388	0.3195	3.3310

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	2.383	1.544

Number of obs: 510, groups: participant_private_id, 170

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.924	0.6409	-6.1226	<0.00001
L2	6.2319	0.9576	6.50749	<0.00001
ASD	1.5686	0.4751	3.30162	0.00096

5.2.Disability spread:

To what extent will college students selectively associate disability cues (i.e., disordered speech, amputation, wheelchair) but not foreignness cues (i.e., foreign accent and garb) with greater dependence?

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Dependence ~ Image_type + (1 | participant_private_id)
Data: help

REML criterion at convergence: 1148.3

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.8634	-0.6953	-0.2664	0.8762	3.0442

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.02775	0.1666
Residual		0.28604	0.5348

Number of obs: 680, groups: participant_private_id, 170

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	1.47058	0.0429	660.5	34.229	<0.00001
Amputee	0.63529	0.0580	506.9	10.951	<0.00001
Wheelchair	1.09411	0.05801	506.9	18.86	<0.00001
Foreign	-0.1352	0.05801	506.9	-2.332	0.02007

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Dependence ~ audio_type + (1 | participant_private_id)
Data: help

REML criterion at convergence: 1028.9

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.2305	-0.7223	0.1516	0.4716	2.1407

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.0892	0.2987
Residual		0.3589	0.5991

Number of obs: 510, groups: participant_private_id, 170

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	1.8235	0.051	469.7	35.516	<0.00001
L2	-0.158	0.064	338.0	-2.444	0.015
ASD	0.3176	0.064	338.0	4.8881	0.000001

To what extent will college students selectively associate reduced physical competence (running ability) with ASD speakers but not L2 speakers?
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Competence ~ audio_type + (1 | participant_private_id)
Data: running

REML criterion at convergence: 714.2

Scaled residuals:
```
    Min      1Q  Median      3Q     Max
-2.29554 -0.13696  0.04177  0.33050  2.77722
```

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.06377	0.2525
Residual		0.18305	0.4278

Number of obs: 510, groups: participant_private_id, 170

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.0235	0.0381	447.2	53.106	<0.00001
L2	-0.0764	0.0464	338.0	-1.6478	0.10031
ASD	-0.1999	0.0464	338.0	-4.3097	0.00002

5.3.Peer Rejection

To what extent will participants express reduced interest in pursuing friendships with foreign or disabled targets (as indicated by speech and/or visual appearance) in comparison to non-foreign, non-disabled targets?

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: response_numeric ~ Image_type + (1 | participant_private_id)
Data: friendship

REML criterion at convergence: 1022.1

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.4601	-0.2745	0.1001	0.4286	3.5595

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.1961	0.4428
Residual		0.1667	0.4083

Number of obs: 680, groups: participant_private_id, 170

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.1058	0.046	360.2	45.58	<0.00001
Amputee	0.1529	0.0442	507.0	3.453	0.0006
Wheelchair	0.0941	0.0442	507.0	2.125	0.034
Foreign	0.0411	0.0442	507.0	0.9298	0.3528

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Friendship ~ audio_type + (1 | participant_private_id)
Data: friendship

REML criterion at convergence: 1002.2

Scaled residuals:
   Min   1Q   Median   3Q   Max
-2.52669 -0.66331 0.01956 0.56525 2.64516

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.2272	0.4767
Residual		0.2693	0.5189

Number of obs: 510, groups: participant_private_id, 170

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	1.9294	0.0540	357.3	35.7019	<0.00001
L2	0.0823	0.0562	338.0	1.46314	0.1443
ASD	-0.123	0.0562	338.0	-2.19472	0.0288

6. Study 1b Aggregate Sample Analysis

6.1.Conceptual differentiation:

To what extent will children selectively associate physical disability cues (i.e., amputation, wheelchair) but not foreignness cues (i.e., foreign garb) with reduced physical competence (running ability)?

Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Competence ~ Image_type * age + (1 | participant_private_id)
Data: run

REML criterion at convergence: 1363.3

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.41133	-0.73295	-0.02556	0.71927	2.55009

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.08749	0.2958
Residual		0.49359	0.7026

Number of obs: 592, groups: participant_private_id, 151

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.555403	0.062645	548.5	40.79148	< 0.0001
Amputee	-1.039057	0.081442	439.0	-12.7582	< 0.0001
Wheelchair	-0.906248	0.081988	438.4	-11.05334	< 0.0001
Foreign	-0.352059	0.081710	438.8	-4.30863	< 0.0001
Age	0.119584	0.062750	548.9	1.90573	0.0572
Amputee x Age	-0.46012	0.081450	440.0	-5.6491	< 0.0001
Wheelchair x Age	-0.36647	0.082275	438.8	-4.4543	< 0.0001
Foreign x Age	-0.22966	0.081724	438.1	-2.8102	0.0052

To what extent will children selectively associate foreignness cues (i.e., foreign accent and garb) but not disability cues (i.e., amputation, wheelchair) with being from another country?

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
Family: binomial (logit)
Formula: prob_select_faraway ~ Image_type * age + (1 | participant_private_id)
Data: location_visual

AIC	BIC	logLik	deviance	df.resid
777.9	817.4	-380.0	759.9	584

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.2076	-0.8251	0.4700	0.7777	2.3988

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.5289	0.7273

Number of obs: 593, groups: participant_private_id, 152

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.2821	0.191055	-1.47704	0.13966
Amputee	0.01570	0.25487	0.061611	0.95087
Wheelchair	0.38512	0.2557	1.505962	0.13207
Foreign	1.19630	0.2682	4.459523	0.000008
Age	-0.6175	0.2065	-2.99061	0.002784
Amputee x Age	0.05848	0.272	0.214405	0.830230
Wheelchair x Age	0.10507	0.271	0.3877	0.698238
Foreign x Age	0.97694	0.284	3.43334	0.000596

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
Family: binomial (logit)
Formula: prob_select_faraway ~ audio_type * age + (1 | participant_private_id)
Data: location_audio

AIC	BIC	logLik	deviance	df.resid
559.5	587.8	-272.8	545.5	414

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.7594	-1.0864	0.6777	0.7885	2.1833

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.002211	0.04702

Number of obs: 421, groups: participant_private_id, 142

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.09903	0.1812	-0.5464	0.5847
L2	0.771923	0.2571	3.0022	0.002
ASD	0.341828	0.2493	1.3708	0.17
Age	-0.81509	0.2100	-3.881	0.0001
L2 x Age	0.906097	0.2785	3.2531	0.0011
ASD x Age	0.582224	0.2704	2.1529	0.0313

6.2.Disability spread:
To what extent will children selectively associate disability cues (i.e., disordered speech, amputation, wheelchair) but not foreignness cues (i.e., foreign accent and garb) with greater dependence?
Linear mixed model fit by REML. t-tests use Satterthwaite's method [lmerModLmerTest]
Formula: Dependence ~ Image_type * age + (1 | participant_private_id)
   Data: help

REML criterion at convergence: 1400.6

Scaled residuals:
    Min   1Q Median   3Q   Max
-2.2826 -0.7604 0.2259 0.7635 2.0426

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.09891	0.3145
Residual		0.52244	0.7228

Number of obs: 592, groups: participant_private_id, 151

Fixed effects:

	Estimate	*Std. Error*	*df*	*t value*	*Pr(>|t|)*
(Intercept)	*1.7308*	*0.06456*	*544.2*	*26.80725*	*<0.00001*
Amputee	*0.7914*	*0.08367*	*440.0*	*9.45918*	*<0.00001*
Wheelchair	*0.7567*	*0.08423*	*439.4*	*8.9842*	*<0.00001*
Foreign	*0.0531*	*0.08406*	*438.9*	*0.631832*	*0.527825*
Age	*-0.2985*	*0.06464*	*544.7*	*-4.6181237*	*0.000005*
Amputee x Age	*0.4034*	*0.08367*	*441.2*	*4.821724*	*0.000002*
Wheelchair x Age	*0.2324*	*0.08452*	*440.0*	*2.749878*	*0.0062*
Foreign x Age	*0.0404*	*0.08406*	*438.5*	*0.48097325*	*0.6307*

Linear mixed model fit by REML. t-tests use Satterthwaite's method [lmerModLmerTest]
Formula: Dependence ~ audio_type * age + (1 | participant_private_id)
  Data: help

REML criterion at convergence: 1041.1

Scaled residuals:
    Min   1Q Median   3Q   Max
-1.83866 -0.69543 -0.03164 0.76430 1.89767

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.2129	0.4614
Residual		0.4887	0.6991

Number of obs: 427, groups: participant_private_id, 148

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	1.88368	0.06993	361.4	26.9342	<0.00001
L2	0.15985	0.08322	281.7	1.920741	0.055
ASD	0.16113	0.08293	281.9	1.942845	0.053
Age	-0.23322	0.06961	362.2	-3.3503	0.0009
L2 x Age	0.11525	0.08338	285.2	1.38218	0.168
ASD x Age	0.02323	0.08303	282.8	0.279823	0.7798

To what extent will children selectively associate reduced physical competence (running ability) with ASD speakers but not L2 speakers?
Linear mixed model fit by REML. t-tests use Satterthwaite's method [lmerModLmerTest]
Formula: Competence ~ audio_type * age + (1 | participant_private_id)
  Data: running

REML criterion at convergence: 1012.6

Scaled residuals:
   Min   1Q Median   3Q   Max
-1.9877 -0.5957 0.3990 0.6744 1.5724

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.1628	0.4035
Residual		0.4875	0.6982

Number of obs: 424, groups: participant_private_id, 145

Fixed effects:

	Estimate	*Std. Error*	*df*	*t value*	*Pr(>|t|)*
(Intercept)	*2.348*	*0.0674*	*373.3*	*34.835*	<0.00001
L2	*-0.0987*	*0.0833*	*278.8*	*-1.185*	*0.237*
ASD	*0.02781*	*0.0828*	*279.8*	*0.3355*	*0.737*
Age	*-0.04479*	*0.067*	*374.4*	*-0.6676*	*0.5047*
L2 x Age	*-0.0403*	*0.0837*	*281.4*	*-0.481*	*0.630*
ASD x Age	*-0.01158*	*0.0827*	*282.9*	*-0.14*	*0.8888*

6.3.Peer Rejection

To what extent will children express reduced interest in pursuing friendships with foreign or disabled targets (as indicated by speech and/or visual appearance) in comparison to non-foreign, non-disabled targets?

Linear mixed model fit by REML. t-tests use Satterthwaite's method [lmerModLmerTest]
Formula: Friendship ~ Image_type * age + (1 | participant_private_id)
  Data: friendship

REML criterion at convergence: 1363.7

Scaled residuals:
    Min   1Q Median   3Q   Max
-2.36645 -0.58798 0.07486 0.60897 2.35155

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.3049	0.5521
Residual		0.3951	0.6286

Number of obs: 593, groups: participant_private_id, 151

Fixed effects:

	Estimate	*Std. Error*	*df*	*t value*	*Pr(>|t|)*
(Intercept)	*2.227291*	*0.068456*	*377.2*	*32.536*	<0.00001
Amputee	*-0.107557*	*0.072933*	*438.3*	*-1.4747*	*0.141*
Wheelchair	*-0.220066*	*0.073165*	*437.8*	*-3.0077*	*0.0027*
Foreign	*-0.229422*	*0.073094*	*438.7*	*-3.1387*	*0.0018*
Age	*0.1142669*	*0.068587*	*377.7*	*1.66601*	*0.0965*
Amputee x Age	*-0.04919*	*0.073093*	*438.9*	*-0.6730*	*0.5012*
Wheelchair x Age	*-0.02548*	*0.073211*	*437.6*	*-0.34804*	*0.7279*
Foreign x Age	*-0.12079*	*0.0731*	*438.4*	*-1.651436*	*0.09936*

Linear mixed model fit by REML. t-tests use Satterthwaite's method [lmerModLmerTest]
Formula: Friendship ~ audio_type * age + (1 | participant_private_id)
  Data: friendship

REML criterion at convergence: 1045.9

Scaled residuals:
    Min   1Q Median   3Q   Max
-1.95614 -0.62746 -0.00949 0.63849 1.80985

Random effects:

Groups	Name	Variance	Std.Dev.
Participants_private_id	(Intercept)	0.2657	0.5154
Residual		0.4767	0.6904

Number of obs: 425, groups: participant_private_id, 143

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.1446	0.0722	335.3	29.68	<0.00001
L2	-0.135	0.0822	279.6	-1.649	0.1002
ASD	-0.0859	0.0820	279.2	-1.047	0.295
Age	0.06017	0.0717	334.2	0.838	0.402
L2 x Age	-0.03546	0.0823	281.2	-0.43	0.667
ASD x Age	-0.044995	0.0819	280.2	-0.5492	0.583

REFERENCES

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., and Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. Behavior Research Methods, 52(1), 388–407. https://doi.org/1 0.3758/s13428-019-01237-x

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258. https://doi.org/10.1037/h0055756

Babik, I., and Gardner, E. S. (2021). Factors affecting the perception of disability: A developmental perspective. *Frontiers in Psychology*, *12*, 702166. https://doi.org/10.3389/fpsyg.2021.702166

Baharloo, R., Vasil, N., Ellwood-Lowe, M. E., and Srinivasan, M. (2023). Children's use of pragmatic inference to learn about the social world. *Developmental Science*, *26*(3), e13333. https://doi.org/10.1111/desc.13333

Baldwin, D. A., Markman, E. M., and Melartin, R. L. (1993). Infants' ability to draw inferences about nonobvious object properties: Evidence from exploratory play. *Child development*, *64*(3), 711-728. https://doi.org/10.1111/j.1467-8624.1993.tb02938.x

Barrett, M. (2007). Children's knowledge, beliefs and feelings about nations and national groups. Hove, England: Psychology Press.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed effects models using lme4. Journal of Statistical Software, 67, 1–48. https://doi.org/10.18637/jss.v067.i01

Bettens, K., Alighieri, C., Bruneel, L., De Meulemeester, L., and Van Lierde, K. (2020). Peer attitudes toward children with cleft (lip and) palate related to speech comprehensibility, hypernasality and articulation. *Journal of Communication Disorders*, *85*, 105991. https://doi.org/10.1016/j.jcomdis.2020.105991

Blood G. W., Blood I. M., and Danhauer J. L. (1977). The hearing-aid "effect." *Hearing Instruments*, 20, 12.

Blood, G. W., and Hyman, M. (1977). Children's perception of nasal resonance. *Journal of Speech and Hearing Disorders*, *42*(3), 446-448. https://doi.org/10.1044/jshd.4203.446

Bogart, K. R., Rosa, N. M., and Slepian, M. L. (2019). Born that way or became that way: Stigma toward congenital versus acquired disability. *Group Processes and Intergroup Relations*, *22*(4), 594-612. https://doi.org/10.1177/1368430218757897

Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American speech*, *82*(1), 32-64. https://doi.org/10.1215/00031283-2007-002

Chapman, R. (2020). The reality of autism: On the metaphysics of disorder and diversity. *Philosophical Psychology*, *33*(6), 799-819. https://doi.org/10.1080/09515089.2020.1751103

Chomsky, N. (1987). *Language and problems of knowledge: The Managua lectures* (Vol. 16). MIT press.

Creel, S. C. (2012). Preschoolers' use of talker information in on-line comprehension. *Child Development*, *83*(6), 2042-2056. https://doi.org/10.1111/j.1467-8624.2012.01816.x

Creel, S. C. (2018). Accent detection and social cognition: Evidence of protracted learning. *Developmental Science*, *21*(2), e12524. https://doi.org/10.1111/desc.12524

Creel, S. C., and Jimenez, S. R. (2012). Differences in talker recognition by preschoolers and adults. *Journal of Experimental Child Psychology*, *113*(4), 487-509. https://doi.org/10.1016/j.jecp.2012.07.007

Cuddy, A. J., Fiske, S. T., and Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in experimental social psychology*, *40*, 61-149. https://doi.org/10.1016/S0065-2601(07)00002-0

Diamond, K. E., and Hestenes, L. L. (1996). Preschool children's conceptions of disabilities: The salience of disability in children's ideas about others. *Topics in Early Childhood Special Education*, *16*(4), 458-475. https://doi.org/10.1177/027112149601600406

Dunn, D. (2015). The social psychology of disability. New York, NY: Oxford University Press.

Eckert, P. (2008). Variation and the indexical field 1. *Journal of sociolinguistics*, *12*(4), 453-476. https://doi.org/10.1111/j.1467-9841.2008.00374.x

Faul, F., Erdfelder, E., Buchner, A., and Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. Infant Behavior and Development, 8(2), 181–195. https://doi.org/10.1016/S0163-6383(85)80005-9

Franck, A. L., Jackson, R. A., Pimentel, J. T., and Greenwood, G. S. (2003). School-age children's perceptions of a person who stutters. *Journal of Fluency Disorders*, *28*(1), 1-15. https://doi.org/10.1016/S0094-730X(03)00002-0

Freeby, N., and Madison, C. L. (1989). Children's Perceptions of Peers with Articulation Disorders. *Child Study Journal*, *19*(2), 133-44.

Gardner, R. C., and Lambert, W. E. (1972). Attitudes and Motivation in Second Language Learning. Rowley, MA: Newbury House.

Garrett, P. (2010). *Attitudes to language*. Cambridge University Press.

Gascoine, L., Higgins, S., and Wall, K. (2017). The assessment of metacognition in children aged 4–16 years: a systematic review. *Review of Education*, *5*(1), 3-57. https://doi.org/10.1002/rev3.3077

Gelman, S. A., and Taylor, M. G. (2014). Gender essentialism in cognitive development. In *Toward a feminist developmental psychology* (pp. 169-190). Routledge.

Gertner, B. L., Rice, M. L., and Hadley, P. A. (1994). Influence of communicative competence on peer preferences in a preschool classroom. *Journal of Speech, Language, and Hearing Research*, *37*(4), 913-923. https://doi.org/10.1044/jshr.3704.913

Ghio, A., Gasquet-Cyrus, M., Roquel, J., and Giovanni, A. (2013). Perceptual interference between regional accent and voice/speech disorders. In *Interspeech 2013 14th Annual Conference of the International Speech Communication Association* (pp. 2138-2142).

Goffman, E. (1963). Stigma: Notes on the management of spoiled identity Englewood Cliffs, NJ: Prentice-Hall.

Goldstone, R. L., and Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69-78. https://doi.org/10.1002/wcs.26

Gordon, E. D., Minnes, P. M., and Holden, R. R. (1990). The structure of attitudes toward persons with a disability, when specific disability and context are considered. *Rehabilitation Psychology*, *35*(2), 79. https://doi.org/10.1037/h0079054

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., and Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in cognitive sciences*, *14*(8), 357-364. https://doi.org/10.1016/j.tics.2010.05.004

Halkias, G., and Diamantopoulos, A. (2020). Universal dimensions of individuals' perception: Revisiting the operationalization of warmth and competence with a mixed-method approach. *International Journal of Research in Marketing*, *37*(4), 714-736. https://doi.org/10.1016/j.ijresmar.2020.02.004

Hall, B. J. C. (1991). Attitudes of fourth and sixth graders toward peers with mild articulation disorders. *Language, Speech, and Hearing Services in Schools*, *22*(1), 334-340. https://doi.org/10.1044/0161-1461.2201.334

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and social psychology review*, *10*(3), 252-264. https://doi.org/10.1207/s15327957pspr1003_4

Heaton, J. B. (1966). Composition through pictures. Essex: Longman.

Hussak, L. J., and Cimpian, A. (2019). “It feels like it’s in your body”: How children in the United States think about nationality. *Journal of Experimental Psychology: General, 148*(7), 1153–1168. https://doi.org/10.1037/xge0000567

Hilton, C. B., Hilton, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., Simson, J., Knox, D., Glowacki, L., Alemu, E., Galbarczyk, A., Jasienska, G., Neff, M. B., Martin, A., Cirelli, L. K., Trehub, S. E., Song, J., Kim, M., and Mehr, S. A. (2022). Acoustic regularities in infant-directed speech and song across cultures. Nature Human Behaviour, 6(11), 1–12. https:// doi.org/10.1038/s41562-022-01410-x

Hirschfeld, L. A. (2003). On a folk theory of society: Children, evolution, and mental representations of social groups. In *Lay Theories and Their Role in the Perception of Social Groups* (pp. 107-117). Psychology Press.

Hirschfeld, L. A., and Gelman, S. A. (1997). What young children think about the relationship between language variation and social difference. *Cognitive Development*, *12*(2), 213-238. https://doi.org/10.1016/S0885-2014(97)90014-9

Horowitz, A. C., and Frank, M. C. (2016). Children's pragmatic inferences as a route for learning about the world. *Child development*, *87*(3), 807-819. https://doi.org/10.1111/cdev.12527

Hrdy, S. B. (2009). *Mothers and others:* The evolutionary origins of mutual understanding. Harvard University Press.

Jones, Z., Yan, Q., Wagner, L., and Clopper, C. G. (2017). The development of dialect classification across the lifespan. *Journal of Phonetics*, *60*, 20-37. https://doi.org/10.1016/j.wocn.2016.11.001

Killen, M., Rutland, A., and Ruck, M. D. (2011). Promoting equity, tolerance, and justice in childhood. *Social Policy Report*, *25*(4), 1-33.

Kinzler, K. D., Shutts, K., and Correll, J. (2010). Priorities in social categories. *European Journal of Social Psychology*, *40*(4), 581-592. https://doi.org/10.1002/ejsp.739

Kinzler, K. D., and DeJesus, J. M. (2013). Children's sociolinguistic evaluations of nice foreigners and mean Americans. *Developmental Psychology*, *49*(4), 655. https://doi.org/10.1037/a0028740

Kinzler, K. D. (2021). Language as a social cue. *Annual review of psychology*, *72*, 241-264. https://doi.org/10/1146/annurev-psych-010418-103034

Kircher, R. (2015). The matched-guise technique. *Research methods in intercultural communication: A practical guide*, 196-211. https://doi.org/10.1002/9781119166283.ch13

Koschorke, M., Evans-Lacko, S., Sartorius, N., and Thornicroft, G. (2017). Stigma in different cultures. *The stigma of mental illness-end of the story?*, 67-82.

Krueger, B. I. (2019). Eligibility and speech sound disorders: Assessment of social impact. *Perspectives of the ASHA Special Interest Groups*, *4*(1), 85-90. https://doi.org/10.1044/2018_PERS-SIG1-2018-0016

Kuczaj, S. and Harbaugh, B. (1982). What children think about the speaking capacities of other persons and things. In S. Kuczaj (Ed.), Language development, Vol. 2, Language. thought and culture. Hillsdale. NJ: Erlbaum.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Labov, W. (1964). Stages in the acquisition of standard English. In R. Shuy (Ed.), Social dialects and language learning (pp. 77–104). Champaign, IL: National Council of Teachers of English.

Lass, N. J., Ruscello, D. M., Bradshaw, K. H., and Blankenship, B. L. (1991). Adolescents' perceptions of normal and voice-disordered children. *Journal of communication disorders*, *24*(4), 267-274. https://doi.org/10.1016/0021-9924(91)90002-Z

Lass, N. J., Ruscello, D. M., Harkins, K. E., and Blankenship, B. L. (1993). A comparative study of adolescents' perceptions of normal-speaking and dysarthric children. *Journal of communication disorders*, *26*(1), 3-12. https://doi.org/10.1016/0021-9924(93)90012-Y

Lee, A., Gibbon, F. E., and Spivey, K. (2017). Children's attitudes toward peers with unintelligible speech associated with cleft lip and/or palate. *The Cleft palate-craniofacial journal*, *54*(3), 262-268. https://doi.org/10.1597/15-088

Lee, K., Quinn, P. C., and Heyman, G. D. (2017). Rethinking the emergence and development of implicit racial bias: A perceptual-social linkage hypothesis. *New perspectives on human development*, 27-46.

Lenth, R. V. (2022). *Emmeans: Estimated marginal means, aka least-squares means* (Version 1.7.2) [Computer software]. https://CRAN.R-project.org/package=emmeans

Leslie, S. J. (2017). The original sin of cognition. *The Journal of Philosophy*, *114*(8), 395-421.

Liberman, Z., Woodward, A. L., Sullivan, K. R., and Kinzler, K. D. (2016). Early emerging system for reasoning about the social nature of food. *Proceedings of the National Academy of Sciences*, *113*(34), 9480-9485. https://doi.org/10.1073/pnas.1605456113

Liberman, Z., Woodward, A. L., and Kinzler, K. D. (2017). Preverbal infants infer third-party social relationships based on language. *Cognitive Science*, *41*, 622-634. https://doi.org/10.1111/cogs.12403

Liesener, J. J., and Mills, J. (1999). An Experimental Study of Disability Spread: Talking to an Adult in a Wheelchair Like a Child 1. *Journal of Applied Social Psychology*, *29*(10), 2083-2092. https://doi.org/10.1111/j.1559-1816.1999.tb02296.x

Lindsay, G., and Strand, S. (2016). Children with language impairment: Prevalence, associated difficulties, and ethnic disproportionality in an English population. In *Frontiers in Education* (Vol. 1, p. 2). Frontiers. https://doi.org/10.3389/feduc.2016.00002

Llanes, À. 2012. The short- and long- term effects of a short study abroad experience: the case of children. System 40: 179-190. https://doi.org/10.1016/j.system.2012.05.003

Loury, G. C. (2005). Racial stigma and its consequences. *Focus*, *24*(1), 1-6.

Loury, G. C. (2021). *The anatomy of racial inequality: With a new preface*. Harvard University Press.

Mandalaywala, T. M., Ranger-Murdock, G., Amodio, D. M., and Rhodes, M. (2019). The nature and consequences of essentialist beliefs about race in early childhood. *Child development*, *90*(4), e437-e453. https://doi.org/10.1111/cdev.13008

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child development*, 986-992. https://doi.org/10.2307/1128350

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child development*, 643-655. https://doi.org/10.2307/1128929

Markman, E. M. (1989). Categorization and naming in children: Problems of induction. MIT Press.

McCoy, S., and Banks, J. (2012). Simply academic? Why children with special educational needs don't like school. *European Journal of Special Needs Education*, *27*(1), 81-97. https://doi.org/10.1080/08856257.2011.640487

Medin, D. (1989). Concepts and conceptual structures. American Psychologist, 45, 1469– 1481. https://doi.org/10.1037/0003-066X.44.12.1469

Mei, C., Reilly, S., Bickerton, M., Mensah, F., Turner, S., Kumaranayagam, D., ... and Morgan, A. T. (2020). Speech in children with cerebral palsy. *Developmental Medicine and Child Neurology*, *62*(12), 1374-1382. https://doi.org/10.1111/dmcn.14592

Miller, P. H., Baxter, S. D., Royer, J. A., Hitchcock, D. B., Smith, A. F., Collins, K. L., ... and Finney, C. J. (2015). Children's social desirability: Effects of test assessment mode. *Personality and individual differences*, *83*, 85-90. https://doi.org/10.1016/j.paid.2015.03.039

Muñoz, C. (Ed.). (2006). *Age and the rate of foreign language learning* (Vol. 19). Multilingual Matters.

Munyi, C. W. (2012). Past and present perceptions towards disability: A historical perspective. *Disability studies quarterly*, *32*(2). doi:10.18061/dsq.v32i2.3197

Nario-Redmond, M. R., Kemerling, A. A., and Silverman, A. (2019). Hostile, benevolent, and ambivalent ableism: Contemporary manifestations. *Journal of Social Issues*, *75*(3), 726-756. https://doi.org/10.1111/josi.12337

Nario-Redmond, M. R. (2024). Contagion Fear: An Evolved Form of Disability Prejudice. In *The Palgrave Encyclopedia of Disability* (pp. 1-6). Cham: Springer Nature Switzerland.

Nejjari, W., Gerritsen, M., Van Hout, R., and Planken, B. (2019). Refinement of the matched-guise technique for the study of the effect of non-native accents compared to native accents. *Lingua*, *219*, 90-105. https://doi.org/10.1016/j.lingua.2018.12.001

Neuberg, S. L., and Cottrell, C. A. (2013). Evolutionary bases of prejudices. In *Evolution and social psychology* (pp. 163-187). Psychology Press.

Nordberg, A., Miniscalco, C., Lohmander, A., and Himmelmann, K. (2013). Speech problems affect more than one in two children with cerebral palsy: Swedish population-based study. *Acta paediatrica*, *102*(2), 161-166. https://doi.org/10.1111/apa.12076

Paquette-Smith, M., Buckler, H., White, K. S., Choi, J., and Johnson, E. K. (2019). The effect of accent exposure on children's sociolinguistic evaluation of peers. *Developmental Psychology*, *55*(4), 809. https://doi.org/10.1037/dev0000659

Parker, L. R., Monteith, M. J., and South, S. C. (2020). Dehumanization, prejudice, and social policy beliefs concerning people with developmental disabilities. *Group Processes and Intergroup Relations*, *23*(2), 262-284. https://doi.org/10.1177/1368430218809891

Peretz-Lange, R. (2021). Why does social essentialism sometimes promote, and other times mitigate, prejudice development? A causal discounting perspective. *Cognitive Development*, *59*, 101085. https://doi.org/10.1016/j.cogdev.2021.101085

Perkins, M. (2007). Pragmatic impairment. Cambridge, UK: Cambridge University Press.

Quine, W.V.O. (1960). Word and object. Cambridge. MA: MIT Press

R Core Team. (2020). *R: A language and environment for statistical computing* (Version 4.1.3) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rakić, T., Steffens, M. C., and Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of personality and social psychology*, *100*(1), 16. https://doi.org/10.1037/a0021522

Ramsey, P. (1987). Young children's thinking about ethnic differences. In J. Phinney and M. Rotheram (Eds.), Children's ethnic socialization: Pluralism and development (pp. 56-72). Beverly Hills, CA: Sage.

Rauterkus, E. P., and Palmer, C. V. (2014). The hearing aid effect in 2013. *Journal of the American Academy of Audiology*, *25*(09), 893-903. https://doi.org/10.3766/jaaa.25.9.10

Redford, M. A., Kapatsinski, V., and Cornell-Fabiano, J. (2018). Lay listener classification and evaluation of typical and atypical children's speech. *Language and Speech*, *61*(2), 277-302. https://doi.org/10.1177/0023830917717758

Reindal, L., Nærland, T., Sund, A. M., Glimsdal, B. A., Andreassen, O. A., and Weidle, B. (2022). The co-occurrence of motor and language impairments in children evaluated for autism spectrum disorder. An explorative study from Norway. *Research in Developmental Disabilities*, *127*, 104256. https://doi.org/10.1016/j.ridd.2022.104256

Rice, M. L., Sell, M. A., and Hadley, P. A. (1991). Social interactions of speech, and language-impaired children. *Journal of Speech, Language, and Hearing Research*, *34*(6), 1299-1307. https://doi.org/10.1044/jshr.3406.1299

Schaller, M., & Neuberg, S. L. (2012). Danger, disease, and the nature of prejudice (s). In *Advances in experimental social psychology* (Vol. 46, pp. 1-54). Academic Press. https://doi.org/10.1016/B978-0-12-394281-4.00001-5

Smith, L. A., and Williams, J. M. (2001). Children's understanding of the physical, cognitive and social consequences of impairments. *Child: Care, Health and Development*, *27*(6), 603-617. https://doi.org/10.1046/j.1365-2214.2001.00236.x

Souza, A. L., Byers-Heinlein, K., and Poulin-Dubois, D. (2013). Bilingual and monolingual children prefer native-accented speakers. *Frontiers in psychology*, *4*, 65278. https://doi.org/10.3389/fpsyg.2013.00953

Spelke, E. S., and Kinzler, K. D. (2007). Core knowledge. *Developmental science*, *10*(1), 89-96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Stanovich, K. E. (2020). Why humans are cognitive misers and what it means for the great rationality debate. In *Routledge handbook of bounded rationality* (pp. 196-206). Routledge.

Tajfel, H., and Turner, J. C. (1979). "An integrative theory of intergroup conflict," in The Social Psychology of Intergroup Relations, eds W. G. Austin and S. Worchel (Monterey, CA: Brooks-Cole), 33–47.

Tavakoli, P., and Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output. *Language learning*, *61*, 37-72. https://doi.org/10.1111/j.1467-9922.2011.00642.x

Wagner, L., Clopper, C. G., and Pate, J. K. (2014). Children's perception of dialect variation. *Journal of child language*, *41*(5), 1062-1084. https://doi.org/10.1017/S0305000913000330

Wagner, L., Greene-Havas, M., and Gillespie, R. (2010). Development in children's comprehension of linguistic register. *Child Development*, *81*(6), 1678-1686. https://doi.org/10.1111/j.1467-8624.2010.01502.x

Weatherhead, D., White, K. S., and Friedman, O. (2016). Where are you from? Preschoolers infer background from accent. *Journal of Experimental Child Psychology*, *143*, 171-178. https://doi.org/10.1016/j.jecp.2015.10.011

Weiner, B. (2005). Motivation from an attribution perspective and the social psychology of perceived competence. *Handbook of competence and motivation*, *2005*, 2-3.

Weiner, B. (2006). ''Social Motivation, Justice, and the Moral Emotions: An Attributional Approach.'' Lawrence Earlbaum, Mahwah, NJ.

Werner, S., Peretz, H., and Roth, D. (2015). Israeli children's attitudes toward children with and without disabilities. *Early Childhood Research Quarterly*, *33*, 98-107. https://doi.org/10.1016/j.ecresq.2015.07.003

West, K. L. (2019). Infant motor development in autism spectrum disorder: A synthesis and meta-analysis. *Child development*, *90*(6), 2053-2070. https://doi.org/10.1111/cdev.13086

Westra, E., and Carruthers, P. (2017). Pragmatic development explains the Theory-of-Mind Scale. *Cognition*, *158*, 165-176. https://doi.org/10.1016/j.cognition.2016.10.021

Wheeler, L. R., and Tharpe, A. M. (2020). Young children's attitudes toward peers who wear hearing aids. *American Journal of Audiology*, *29*(2), 110-119. https://doi.org/10.1044/2019_AJA-19-00082

Wright, B. A. (1964). Spread in adjustment to disability. *Bulletin of the Menninger Clinic*, *28*(4), 198-208.

Wright, B. A. (1983). Physical disability: A psychosocial approach. New York, NY: Harper and Rowe.