

A combined sociolinguistic and experimental phonetic approach  
to loanword variation and adaptation

by

Zachary Jagers

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

New York University

September, 2018

---

Renée Blake

---

Lisa Davidson

## Acknowledgements

Thank *you*, the reader of this work, and all those invested in thinking critically about language, how it works, and what it can reflect about the groupings and even the ideologies in society.

To my co-chairs, thank you for your encouragement of my combined interests. Discussions with each of you over the years pushed me to refine my research while always feeling like it was still my program that I was shaping: the best kind of academic mentorship. Lisa Davidson, the chair I knew I would work with from the start, your work inspired my own and your resolutely empirical approach will continue to influence the way I think about language and the methods of studying it. I'm grateful for your guidance, frankness, around-the-clock responsiveness, and feedback both big-picture and small-scale. Renée Blake, the chair I had no idea I would work with when I started, I'm so glad we connected early in my program. You helped me develop a better research program by continually drawing my attention to the interplay of identities, ideologies, and biases. You helped me develop into a better researcher by pushing me to meditate on how these are also at play within myself, to acknowledge their potential influence in all facets of the academic process, and to check such influence. The energy and passion that you each bring to your research, teaching, and advisement kept me going. I hope to pass these on to others in whatever my next steps may be.

I'm also grateful to the other members of my committee. Greg Guy, from the beginning you have facilitated great discussions and raised sharp questions which honed my linguistic thinking skills. Laurel MacKenzie, your well-prepared and effective teaching and writing serve as an exemplar that I will always strive to emulate. Yoonjung Kang, your research has served as a model and motivation for my own, and I'm so grateful for your willingness to join my committee as an external member. Thank you for your insightful feedback and engaging, collegial discussion by writing and at conferences. I hope to stay in touch. And thank you to other NYU faculty members, like John Singler, Maria Gouskova, Frans Adriaans, Adam Buchwald, Gillian Gallagher, Stephanie Harves, and Anna Szabolcsi, who advised earlier

projects of mine, were model instructors to observe and TA for, and provided stimulating coursework allowing me to take sometimes big leaps to connect my own interests to it.

There were so many others who were crucial to this dissertation coming to fruition. Thank you to all who lent me your time and energy essential to this project by letting me use your voice as a recording for many strangers to hear (or even your expert illustrations for them to see!), by serving as a participant for me to analyze, or by helping me collect and analyze data for some parts of the project. I'm extremely grateful to those in Bloomington who were so generous with their time, resources, and space to host me during data collection: Ken de Jong who helped immensely with logistical setup and gave me practically free reign of the lab's recording facility, Phillip Weirich who provided a helping hand in many ways and a good time with fellow linguists, and Greer Gerni, the fortuitous connection and instant friend completely unforeseen yet readily providing a home away from home and a social scientist's dream of recruitment assistance.

Beyond the things I learned and studied, my time at NYU was transformative because of the many great people who were a part of it. Becky Laturus, it felt so rare and precious that even in NYC we could be both such great friends and physical neighbors. I'm so glad for the many moments we could share together: good food, fun times, immense support, and just stellar company in grad school and life more generally. Suzy Ahn, I was so happy whenever you and I were both in the office. Our simultaneously casual and shop-talk yet also deep and vulnerable conversations together, drinking bubble tea or coffee while strolling through the park, were the times I was most glad to have come to campus. I can't wait for the next, wherever it may be. Nicole Holliday, frientor and first co-author, thank you for all of your modeling, encouragement, pushing, and jibing. Our conversations would always leave me feeling like there's so much more that I can do (and that I want to do).

I'm grateful for the cohort who started the program with me—Becky, Dan Duncan, Isaac Bleaman, Adina Williams, Songhee Kim, Yohei Oseki, Jeff Lin. Our regular check-in lunches early on in the program and our similar passing-by conversations later on were an

instant reminder that I'm not alone. And thank you to those who were already there—Suzy, Nicole, Emily Nguyen, Sean Martin, Nate LaFave, Dani Szeredi, Carina Bauman, Allison Shapp, James Whang, Dunja Veselinović—helping me feel welcome and motivated from the start. I'm also glad to have made valuable connections with those who entered later—Natalie Povilonis de Vilchez, Mary Robinson, Maddie Gilbert, Yining Nie, Ildi Szabó, Sarah Phillips, Sheng-Fu Wang. From the exchange of NYC living tips to commiserations about the PhD student experience, from shared research advice and tools to enlightening theoretical discussion, you made me feel well-situated in an affirming subcommunity within the broader community of academia that can sometimes feel so different.

There were many who fostered my inclination toward linguistics at its incipient stages and along the way. Thank you, Janet Leak-Garcia, for all your friendship and mentorship having just finished the PhD experience right as I was entering it. Jennifer Nycz, my valued advisor of my Master's and model of the kind of linguist I knew I wanted to become, with that aim in mind it was hard deciding whether to stay with you at Georgetown or follow your footsteps through NYU. I'm also glad I got the chance to work with and learn from other Georgetown faculty like Lisa Zsiga, Lourdes Ortega, and Natalie Schilling. I'm thankful for Lourdes Morales-Gudmundsson who advised me through undergrad and supported my interest in cross-linguistic analysis, academic self-direction, and spending a ridiculously large proportion of my undergrad years doing studies abroad all over the world. And the conversation that started it all, with Stephanie Brooks one day in high school after class. (I can't remember whether it was after Spanish class or French class—a clue already.) You said something along the lines of “That question sounds like the kind of question a linguist would ask.” Thank you for that spark, and to all of those mentioned above who helped fan this flame.

Thank you also to those who necessarily reminded me that linguistics is not my entire life's worth. I'm so glad I found Khorikos as early as I did in my time in NYC. Singing with you was always a healing refuge from grad school-induced intellectual fatigue, allowing

me to stretch and connect with my soul more purely for a change. I'm also grateful for the chances I got to sing with (also now Dr.) Nick Zork and the AH crew.

Beyond NYC, I'm grateful for the many invaluable friends who have remained close in other ways. We had fun get-togethers and refreshing conversations when we had the fortune to be in the same space. At the same time, I so often felt a network of supportive confidants, reminding me how true friendship can be empowering and enlivening even across great distances.

I am extremely fortunate to have family who, even if not knowing exactly what I was doing, why I was doing it, or what it might pragmatically buy me, have always encouraged my pursuits and taken care of me in many different ways and at many different times along the way.

Ben Yancer, my closest family and best friend. You have provided unwavering emotional support, been an insightful sounding board, made and shared so many sweet memories, and energized me when I was drained. You've taken on so much un-cited, unpaid labor that I can never appreciate enough but will never stop trying to. I am privileged, because I don't see how I could have done this all without you.

## Abstract

When two languages are in contact, one of the most common and likely things that will happen is that one language will borrow a word from the other (Haugen 1950; Thomason 2001; Sankoff 2002). Of course, since they are not the same language, the borrowing language will likely have to adapt the word's sound form to some degree, therefore not resulting in an exact replication of the source form. The sound system of the borrowing language will dictate this in large part: e.g., the word might contain sounds that the borrowing language doesn't have (or that the borrowing language at least doesn't pronounce in the exact same way as the source language), so the most similar sound available to the borrowing language is used in its stead. Sometimes, though, the end result is not so clear-cut and loanwords can become variable between different pronunciations. In some cases, this can be attributable to the intensity of contact between the two languages and the speaker's degree of familiarity with the source language, where a speaker might use a more source-like pronunciation by maintaining sounds that otherwise wouldn't be considered allowed in the borrowing language (Poplack et al. 1988). But, in other cases, a loanword might vary between pronunciations that more or less closely resemble a loanword's source form *in spite* of each pronunciation being readily available in the borrowing language's native sound system.

It is the latter kind of variation that this dissertation is concerned with. The variation of loanwords in American English between more vs. less source-like pronunciations is examined and tested for how social factors even further removed from intensity of language contact might explain it. It is hypothesized that speakers' relation to or attitude toward the source might predict whether they are more strongly inclined to use a more source-like pronunciation, in line with sociolinguistic theories and evidence that speakers are more likely to resemble each other in their language depending on how much they identify with and want to resemble each other (Labov 1972; Bell 1984; Le Page and Tabouret-Keller 1985; Giles et al. 1991; Milroy and Milroy 1992; Eckert 2004). Furthermore, it is considered that the social relation of interest might not only be the singular language contact situation particular

to a loanword's specific source language. Instead, loanwords as a broad class may pattern together, similarly reflecting something about the social relation of the speaker with respect to foreign languages and/or people as a broader group.

The backdrop to the pursuit at hand is Hall-Lew et al.'s (2010) previous study of the variable pronunciation of *Iraq* in the speech of US politicians, which they observe to pattern with speakers' political identity—Republicans preferring the less source-like [aɪ'ɹæk] pronunciation and Democrats preferring the more source-like [ɪ'ɹɑk] pronunciation. Instead, in Chapter 2, it is hypothesized and empirically corroborated that this political variation falls out from and reflects the attitudes and ideologies that also pattern with political identity in the US political zeitgeist. It is observed that, yes, even amongst non-politicians this variation patterns with political identity; but, other factors like the speaker's source-directed attitude and alignment with a globalist/nationalist ideology are identified as better predictors when accounted for alongside political identity. It is especially the latter factor that is the strongest predictor: A speaker who is more globalist-aligning is more likely to use a more source-like pronunciation. Furthermore, this pattern holds across loanwords of various kinds (not just placenames), sound variables, and sources, reflecting that this ideological alignment regards the broader world and the speaker's relationship to it.

Results do, however, suggest that some variation can still be explained by political identity, which is interpreted to suggest that loanword variation has gained a second-order (Silverstein 2003; Eckert 2008) political indexicality. In Chapter 3, people's perceptions of loanword variation are examined to shed further light on its indexicality. Similar to how it patterns in speech production, the use of more source-like pronunciations is perceptually indexed as more globalist-aligning and globally oriented, as well as being associated with prestige and linguistic security (Labov 1966). While these indexations are apparent in both implicit, matched-guise testing (Lambert et al. 1960) and an explicit, metalinguistic questionnaire, a significant political indexation only surfaces when examined in an explicit method. This is interpreted to corroborate that the political indexation of loanword varia-

tion is second-order, with the difference between implicit perceptual activation vs. explicit commentary suggesting a layered indexicality parallel to how loanword variation is observed to pattern in speech production.

Finally, Chapter 4 cycles back to consider how such variation arises in the first place. In a phonetic imitation experiment (Pardo 2006; Babel 2010; Yu et al. 2013) designed to simulate initial loanword adaptation, participants hear nonce words framed as loanwords, manipulated between phonetic exposure forms but in which both variants are licit and attested in the borrowing language sound system. Results suggest that the variation of established loanwords parallels the adaptation of new loanwords: Those who exhibit closer replication of the exposure forms of new loanwords they've heard are those who hold a more positive attitude toward the source (supporting Weinreich's [1968, 27] hypothesis) and those who are more globalist-aligning. This suggests that loanword variation may arise as a result of socially mediated loanword adaptation. Chapter 5, based on these findings, provides a sketch of how the path of loanword adaptation is socially mediated with the potential of resulting in the variation of even well-established loanwords between more and less source-like variants. Further discussion encourages the field to consider and test the effects loanword status, in and of itself, on sociolinguistic variation (i.e., the co-variation of loanwords across diverse sources) and sound processing, as well as its interaction with social factors like those considered here and beyond.



# Contents

Acknowledgements . . . . .	ii
Abstract . . . . .	vi
List of Figures . . . . .	xii
List of Tables . . . . .	xiv
List of Appendices . . . . .	xv
<b>1 Loanwords: Borrowing, adaptation, and variation</b>	<b>1</b>
<hr/>	
1.1 Introduction . . . . .	1
1.2 Loanword borrowing and adaptation . . . . .	2
1.2.1 Borrowing . . . . .	2
1.2.2 Fitting in: Loanword adaptation . . . . .	3
1.2.3 Getting special treatment: Loanword <i>status</i> as an influence . . . . .	8
1.3 Loanword variation . . . . .	10
1.3.1 A not-quite-arbitrary variable . . . . .	10
1.3.2 Relevant social factors . . . . .	14
1.3.3 Variation as a result of socially mediated adaptation . . . . .	19
1.4 Roadmap . . . . .	22
<b>2 Variation of established loanwords</b>	<b>25</b>
<hr/>	
2.1 Introduction . . . . .	25
2.2 Potential roots to political loanword variation . . . . .	27
2.2.1 Source-directed attitude . . . . .	29
2.2.2 Language contact ideology . . . . .	31
2.2.3 Globalist/Nationalist ideological alignment . . . . .	35
2.3 Empirical analysis . . . . .	41
2.3.1 Study 1: Random sampling questionnaire . . . . .	41
2.3.1.1 Methods . . . . .	42

2.3.1.2	Results . . . . .	43
2.3.2	Study 2: Directly measuring and comparing factors of interest . . . . .	49
2.3.2.1	Methods . . . . .	50
2.3.2.2	Statistical analysis . . . . .	55
2.3.2.3	Results . . . . .	57
2.3.2.4	Interim discussion . . . . .	63
2.3.3	Study 3: Increasing political identity representation; controlling region . . . . .	66
2.3.3.1	Methods . . . . .	66
2.3.3.2	Results . . . . .	68
2.4	Discussion and Conclusions . . . . .	78
<b>3</b>	<b>Perceptual indexation and evaluation of loanword variation</b>	<b>85</b>
<hr/>		
3.1	Introduction: Considering perception . . . . .	85
3.2	Methods . . . . .	91
3.3	Results . . . . .	98
3.3.1	Matched-guise . . . . .	98
3.3.2	Metalinguistic subjective evaluation . . . . .	101
3.3.3	Considering factors of the participant . . . . .	103
3.3.3.1	Metalinguistic ratings by factors of the participant . . . . .	104
3.3.3.2	Matched-guise ratings by factors of the participant . . . . .	106
3.4	Discussion and Conclusions . . . . .	107
<b>4</b>	<b>Loanword adaptation as socially mediated sound imitation</b>	<b>118</b>
<hr/>		
4.1	Introduction . . . . .	118
4.1.1	Socially mediated phonetic imitation . . . . .	119
4.1.2	Loanword variation as mediated source form preservation . . . . .	123
4.2	Methods . . . . .	127
4.2.1	Stimuli . . . . .	127

4.2.1.1	Phonetic variables . . . . .	127
4.2.1.2	Framing . . . . .	128
4.2.1.3	Preparation . . . . .	131
4.2.2	Procedure . . . . .	133
4.3	Results . . . . .	136
4.3.1	Categorical imitation results . . . . .	137
4.3.2	Phonetic imitation . . . . .	146
4.3.2.1	Acoustic analysis and imitation scoring . . . . .	147
4.3.2.2	Phonetic imitation results . . . . .	154
4.4	Discussion and Conclusions . . . . .	160
<b>5</b>	<b>Conclusion</b>	<b>168</b>
<hr/>		
5.1	Summary of findings . . . . .	168
5.2	Theoretical implications . . . . .	171
5.2.1	The socially mediated path of adaptation . . . . .	171
5.2.2	Indexicality . . . . .	176
5.3	Limitations and future directions . . . . .	181
5.3.1	Expanding to other contexts . . . . .	181
5.3.2	Processing considerations . . . . .	184
5.3.3	Loanword status as an influence . . . . .	187
	Appendices . . . . .	192
	References . . . . .	225

## List of Figures

1	Random sampling questionnaire: raw results . . . . .	44
2	Random sampling questionnaire: Rbrul results . . . . .	45
3	Cross-tabulations: political identity $\times$ word . . . . .	47
4	Political identity as predictor of $\{Iran, Iraq\}$ variation . . . . .	58
5	Middle East-directed attitudes as predictors of $\{Iran, Iraq\}$ variation . . . . .	59
6	Ideologies as predictors of $\{Iran, Iraq\}$ variation . . . . .	60
7	Political identity distribution of participants . . . . .	69
8	Political identity as predictor of $\{Iran, Iraq\}$ variation . . . . .	71
9	Globalist/Nationalist alignment as predictor of $\{Iran, Iraq\}$ variation . . . . .	72
10	Prescriptivism as predictor of $\{Iran, Iraq\}$ variation . . . . .	74
11	Political identity, globalist/nationalist alignment as predictors of general loanword variation . . . . .	75
12	Mobility, prescriptivism as predictors of general loanword variation . . . . .	77
13	Journalistic reports on Obama’s pronunciations of foreign words and names . . . . .	86
14	Matched-guise results . . . . .	99
15	Explicit metalinguistic rating results . . . . .	101
16	Distribution of participants: political identity $\times$ $Ir[a\sim\ae]q$ pronunciation . . . . .	103
17	Sample illustrations including orthographic labeling . . . . .	131
18	Categorical imitation: word $\times$ framing . . . . .	139
19	Categorical imitation: globalist/nationalist alignment . . . . .	141
20	Categorical imitation: place-directed attitude . . . . .	142
21	Categorical imitation: language contact ideology . . . . .	144
22	Categorical imitation: political identity . . . . .	145
23	Baseline transition earliness of $[j\partial]$ vs. $[i\partial]$ . . . . .	148
24	General imitation of exposure form: GLIDING variable . . . . .	149
25	Baseline $[\partial]$ vs. $[\epsilon]$ . . . . .	151

26	General imitation of exposure form: VOWEL REDUCTION variable . . . . .	152
27	Stimulus [ə] vs. [ɛ] in comparison to participant baselines . . . . .	153
28	F2 by [ə] vs. [ɛ] exposure: VOWEL REDUCTION variable . . . . .	154
29	General phonetic imitation: globalist/nationalist alignment . . . . .	155
30	General phonetic imitation: other social factors of the individual . . . . .	156
31	Phonetic imitation: globalist/nationalist alignment, across manipulated variables . . . . .	157
32	Phonetic imitation of gliding: place-directed IAT . . . . .	159

## List of Tables

1	Variable words of interest: more source-like $\sim$ less source-like . . . . .	51
2	Example Likert agreement statements . . . . .	53
3	Lone-predictor statistical results . . . . .	61
4	Collinearity of social factors . . . . .	62
5	Step-up comparison results . . . . .	62
6	Variable words of interest: more source-like $\sim$ less source-like . . . . .	67
7	Likert agreement statements: prescriptivism . . . . .	68
8	Collinearity with political identity . . . . .	70
9	Collinearity of social factors . . . . .	72
10	Lone-predictor statistical results: $\{Iran, Iraq\}$ variation . . . . .	73
11	Step-up comparison results: $\{Iran, Iraq\}$ variation . . . . .	73
12	Lone-predictor statistical results: All target loans . . . . .	76
13	Step-up comparison results: All target loans . . . . .	76
14	Word-specific intercepts and political identity slopes . . . . .	78
15	Variable placenames of interest . . . . .	93
16	Report reactions . . . . .	95
17	Explicit Likert ratings . . . . .	97
18	Matched-guise statistics . . . . .	99
19	Explicit metalinguistic rating statistics . . . . .	102
20	Nonce stimuli . . . . .	130
21	Key words elicited for baseline measurements . . . . .	134
22	Example procedure for two (hypothetical) participants . . . . .	135
23	Step-up comparison results: Categorical imitation . . . . .	140
24	Step-up comparison results: Phonetic imitation (gliding) . . . . .	158

## List of Appendices

A	Likert questionnaire . . . . .	192
B	Step-up modeling, Study 2 . . . . .	195
C	Step-up modeling, Study 3 . . . . .	199
D	Matched-guise news reports . . . . .	205
E	Modeling effects of individual factors on sociolinguistic perception . . . . .	206
F	Stories and sequels . . . . .	210
G	Step-up modeling, Nonce word imitation . . . . .	219

# Chapter 1

## Loanwords: Borrowing, adaptation, and variation

### 1.1 Introduction

Loanword borrowing is one of the most frequent and likely outcomes of even light situations of language contact (Haugen 1950; Thomason and Kaufman 1988; Thomason 2001; Sankoff 2002). And, given that a loanword includes (by definition) the imitation of the sound form of another language, its analysis calls for attention to both the sound system and the social context. The sound system of the borrowing language, and how its speakers therefore process the borrowed sound form, can influence the outcomes of this imitation—how closely the borrowing resembles the way the word is pronounced in its source language or how strongly it is adapted to fit in with the borrowing language. But, the imitation of a linguistic form is not solely mediated by grammatical mechanisms; it is also mediated by the social context. Distinct languages, language varieties, or variants of particular language features can arise or merge due to how much different groups interact with each other or, even when interacting with each other, how much they identify with and want to resemble each other (Labov 1972; Bell 1984; Le Page and Tabouret-Keller 1985; Cavalli-Sforza et al. 1988; Giles et al. 1991; Milroy and Milroy 1992; Eckert 2004). Given this understanding, social factors may also influence how (or to what degree) loanwords from one language are adapted when borrowed into another language. The study of loanword adaptation and variation can therefore enlighten our understanding of the sound system, the social context, and how they interact with each other.



## 1.2 Loanword borrowing and adaptation

### 1.2.1 Borrowing

Contact between groups can lead to them sharing linguistic features that they might not otherwise have had in common. In cases of contact between speakers of different languages, this usually comes about via borrowing, where speakers of one language (the borrowing language) will incorporate a feature of another language (the source language). Sankoff (2002, 643) describes this process to most commonly proceed as follows: Native speakers of the borrowing language who also are bilingual with the language they are in contact with “find themselves introducing second-language lexical items into conversation with fellow bilinguals in their original first language.” These ‘nonce borrowings’ (as termed by Weinreich 1968, 47) “seem to constitute the thin end of the wedge in various types of subsequent linguistic change” (continues Sankoff 2002, 643), with the possibility of becoming more established loanwords used by the wider community of speakers of the borrowing language (see also Poplack and Sankoff 1984; van Coetsem 2000; Myers-Scotton 2002) which may carry along with them “not only phonological baggage, but often [...] morphological and syntactic baggage as well.” This entails that lexical borrowings are the most likely outcomes of language contact, which may also lead to further phonological, morphological, and syntactic change.

This resonates with other approaches to borrowing that also suggest lexical borrowing to be the most common manifestation of language contact, though treating this more as a scale of likelihood than with such a directly serial, implicational nature (e.g., Thomason and Kaufman 1988; Thomason 2001). For example, there are cases of language contact that have been identified to bring about the borrowing of semantic features or meanings but without their associated surface forms, which could be considered ‘semantic borrowings’, ‘loan translations’, or ‘calques’ (Haugen 1950). This dissertation is not intended to speak to this debate regarding the relative likelihoods or interdependencies of different kinds of outcomes of language contact, nor is it intended to speak to the debate of what constitutes

a ‘nonce borrowing’/‘code-switching’ or ‘loanword’ or how necessary or possible it is for such a distinction to be formalized (see Myers-Scotton 1993; Haspelmath 2009; Poplack and Dion 2012). However, it should be noted upfront that the subject of study here is lexical borrowings, henceforth referred to as ‘loanwords’: word-level form+meaning pairings used in one language that resemble and are identifiably derived from form+meaning pairings in other languages by means other than shared inheritance from a mutual ancestor language. The ‘form’ part of this equation is crucial to the study at hand, given that loanwords of interest in English are analyzed for how closely their forms in the borrowing language resemble their forms in their respective source languages, as well as how there is variation in this degree of resemblance. And, it is notable that even in this situation of light language contact (on the English end, given that English is a highly dominant contact language) observations will suggest that loanwords are indeed present and meaningfully variable in their degree of source form replication.

### **1.2.2 Fitting in: Loanword adaptation**

As discussed above, sometimes the borrowing of loanwords can lead to the borrowing of certain sounds or an increased allowance for certain phonotactic patterns within them which were not previously considered available in the borrowing language’s sound system (e.g., Nurse 1985; van Coetsem 1988; Kay 1995; Nagy 1996; Thomason 2001; Sankoff 2002). For example, while [f] and [v] were considered allophones of the same phoneme ([v] appearing as a result of intervocalic voicing) in Old English (Mitchell and Robinson 2011, 15), they are considered to have become distinct phonemes in Middle English due to the borrowing of a high volume of French loanwords during the Norman Conquest in which the two sounds remained distinct in positions that would otherwise neutralize the distinction in Old English phonology (Smith 2009a, 23): e.g., *profit*, *refuge*, *veal*, *voice*. However, more often, a loanword is adapted in its form to comply with the sound system of the borrowing language. Itô and Mester (1999) also analyze how the phonology of Japanese was significantly impacted

by extensive borrowing, as further discussed below (§1.2.3). Kang (2011) provides a review of how the adaptation of a loanword's form can be influenced by factors like the borrowing language's sound inventory, phonotactics, and even phonetics.

Sounds in the source form of a loanword that are not members of the borrowing language's sound inventory are often changed to those that are pre-existing members. Of course, given that these are still lexical borrowings and that they therefore include the borrowing of the source form, rather than just its paired meaning, the selection of substitute sounds is not random or arbitrary: Sounds that are as similar as possible to those of the source form are preferred. However, complications arise when considering how sounds are evaluated as optimally similar and therefore selected as substitutes. In some cases, it appears to be the phonological feature system of the borrowing language that dictates this: A sound is preferred if it matches the source sound along the most (or most important) features that carry a contrastive load, as determined by the borrowing language (e.g., Paradis and LaCharité 1997; Clements 2001; Herd 2005; Ito et al. 2006; Drescher 2009; Arsenault 2009). For example, Lombardi (2003) examines a case of 'differential substitution' (Weinberger 1997) in which different languages choose different substitution segments in the adaptation of the same source form segment. Lombardi discusses how, amongst some languages which have both the phonemes /s, t/ in their inventories, some adapt the English interdental fricative /θ/ to /s/ while others adapt it to /t/. Japanese exhibits the /θ/ → /s/ adaptation: e.g., /suɾiɾu/, from *thrill* [θ.ɹɪl] (Kay 1995, 69). Thai, on the other hand exhibits the /θ/ → /t/ adaptation: e.g., /fútpá:t/, from *foot* [fʊt] + *path* [pæθ] (Nacaskul 1979, 158). Lombardi suggests that a difference in these languages' phonological systems explains this phenomenon: Older Japanese words exhibit certain alternations suggesting that the language places a special importance on preserving the manner of articulation of a consonant, leading it to prefer maintaining the fricative manner of the source sound and choose /s/. Thai, on the other hand, does not appear to rank the preservation of a consonant's manner of articulation as high, leading it to resort to /t/, the less inherently marked of the two options.

In other cases, it appears that phonetic similarity is the driving force behind what sound is selected (e.g., Brannen 2002; Steriade 2009; Hsieh et al. 2009). Lin (2008) demonstrates a case where adaptation clearly contradicts what would have been predicted by considering features alone: The vowels in English-to-Mandarin loanwords show a faithfulness to their backness that can override faithfulness to their height, even in contexts where native Mandarin vowels would be neutralized along the feature of frontness/backness while still contrasting by height. For example, the adaptation of *Jackson* [dʒæksən] is [tɕjek<sup>h</sup>ɣsən] (Lin 2008, 370), where the frontness of the source [æ] vowel is preserved and the height is changed in spite of the fact that, in native Mandarin phonology, non-high vowels contrast *only* for height and not for frontness/backness. As Kang (2011) suggests, it is likely that both phonological categorization and phonetic characterization can play a role, seeing that there are single contact situations in which loanword adaptations seem to be influenced by both factors (e.g., Adler 2006; Kenstowicz and Suchato 2006; Rose and Demuth 2006; Chang 2012).

Similarly, loanwords are usually adapted to forms that comply with the phonotactics of the borrowing language as well. For example, a language that does not allow tautosyllabic consonant clusters or word-final consonants may borrow a word containing such a sound form from a source language that does allow them. There is a wealth of studies across many languages finding that a vowel is often inserted either between or before two consonants adjacent in the source form so as to break them up between two syllables, or after a consonant at the end of a word that is not an allowable coda type, therefore resulting in a sound sequence that complies with the borrowing language’s phonotactics (Singh 1985; Hafez 1996; Davidson and Noyer 1997; Katayama 1998; Gouskova 2001; Rose and Demuth 2006; Uffmann 2006; Chang 2012; i.a.). For example, Ulrich (1997) observes that Lama exhibits both types of vowel insertion: The Lama adaptation of English *truck* [tɹʌk] → [teruku] (428) inserts a vowel between the first two consonants [tɹ] since Lama does not allow complex syllable onsets, and it inserts a vowel after the final consonant [k] since the only consonants allowed

as syllable codas are [n, m, l, r, w]. This preference for inserting new material, rather than deleting any (which could also result in phonotactic compliance), is in line with Paradis and LaCharité's (1997) Preservation Principle, suggesting that there is a preference for preserving as much material from the source form as possible. However, there are some cases where sound deletion is observed in adaptation (e.g., Karttunen 1977). For a cluster, this often depends on the types of sounds that compose the cluster (e.g., Silverman 1992; Fleischhacker 2001; Miao 2006; Shinohara 2006) or its position in the word (Chang 2012). We also observe deletion of word-final singleton consonants in borrowing languages that do not allow codas, especially those that also place strong limitations on the number of syllables a word can have (e.g., Silverman 1992; Yip 1993). Other factors of the borrowing language sound system, besides sound inventory and phonotactics, are also observed to influence loanword adaptation, such as prosody (e.g., Kubozono 2006; Yip 2006; Davis et al. 2012; de Jong and Cho 2012; Ito 2014).

Amidst the analysis of what outcomes in loanword adaptation are most frequently observed, there is debate about what processes lead to these outcomes. One major question is whether adaptation takes place during perception or production. Some argue that adaptation takes place during perception, mapping the source form's acoustic realization to its closest corresponding representation available in the borrowing language's sound system (Silverman 1992; Dupoux et al. 1999; Peperkamp et al. 2008; Boersma and Hamann 2009; Kim 2009; Dupoux et al. 2011), with varying proposals considering how certain mechanisms and/or degrees of phonetic detail may influence how the perceived source form is represented and stored. Others contend that adaptation takes place during production, arguing that the initial borrower is a bilingual familiar with the source language (Paradis and LaCharité 1997; Sankoff 2002) and, because of that, the starting/underlying representation is that of the source language which is run through the sound system of the borrowing language when uttered (Hyman 1970; Lovins 1975; Paradis and LaCharité 1997; Shinohara 2004; LaCharité and Paradis 2005). Kang (2011) once again suggests that we should avoid such a

stark, “either... or...” debate, outlining a wealth of studies in which the phonetic realization seems to guide adaptation and override feature preservation, while many other studies (see also LaCharité and Paradis 2005) suggest adaptation to result from the production and preservation of features in spite of other possible realizations being more accurate replications of the source form at the phonetic level.

This dissertation is not intended to speak between these different stances. But, given that they each appear to have support in different cases, this suggests many potential loci of variation between individuals and communities in the way they process loanwords. And, many studies—even those more squarely situated in phonology rather than sociolinguistics—recognize how factors of the context may mediate which kinds of processes more strongly influence any given case of loanword adaptation, such as the individual’s familiarity with the source language (Kang 2010; de Jong and Cho 2012; Ito 2014; Kang et al. 2016), the synchronic point of a language’s history (especially as it regards the relative strength of contact with different languages) during which borrowing takes place (Steinbergs 1985; Nádasdy 1989; Kang 1996, Itô and Mester 1999), or the medium of transmission (spoken vs. orthographic: Smith 2006; Vendelin and Peperkamp 2006; Miao 2006).

A common thread amongst the above-cited studies is phonological non-nativeness. The usual motivation behind the analysis of loanword adaptation is the hope that it can shed light on the linguistic system by observing how forms and features novel to the system of the borrowing language are treated when processed through it. Discrepancies between the source and borrowing languages’ sound systems clearly play a role. The current study, however, acknowledges these effects of the sound system as a strong motivation to either account for or control for them if intending to rigorously analyze how the forms of loanwords are influenced by social and contextual factors. Otherwise, observations could conflate the effects of social factors with the clearly apparent effects of phonological non-nativeness: i.e., One speaker or social group might be more likely than another to use more source-like pronunciations of loanwords from a particular source; but, before conjecturing a social reason, we might first

consider how any differences in the sound systems of the different speakers or social groups could explain the difference observed. The following section will further motivate why we might consider loanword status, in and of itself (and possibly regardless of phonological non-nativeness), to play a role in the sound system. Then (§1.3), we will consider how this factor may not only be meaningful to sound processing but to sociolinguistic variation as well.

### 1.2.3 Getting special treatment: Loanword *status* as an influence

As discussed above, loanwords are usually adapted to a form that complies with the phonological system of the borrowing language. There are cases, though, where loanword adaptation does not behave as expected. In many such cases, loanwords are kept in a form more faithful to the source form than would be expected, preserving sounds that are not members of the borrowing language inventory or sound sequences that the borrowing language's phonotactics would otherwise ban. Analysts suggest that the lexicon is sensitive to whether a word is indexed as foreign or not, with words indexed as more foreign being 'peripheral' and not as strongly adhering to the rules or constraints of the sound system that are apparent when examining the sound forms of words native to the borrowing language (e.g., Kiparsky 1968; Saciuk 1969; Holden 1976; Steinbergs 1985; Itô and Mester 1999).

As a simpler example, Davidson and Noyer (1997) observe that Spanish-to-Huave loanwords can retain the stress placement of their Spanish source form in spite of this conflicting with the stress placement constraints adhered to by native Huave words. As a more elaborate example, Itô and Mester (1999) discuss how phonotactic differences between synchronic groups of words in Japanese can be explained by their relation to diachronic epochs of language contact and borrowing and the impact of these epochs on Japanese phonology. Older (Yamato) words exhibit a constraint against word-internal sequences of a nasal consonant followed by a voiceless obstruent. However, extensive borrowing from Chinese (leading to a stratum of the lexicon referred to as Sino-Japanese vocabulary) carried loanwords with source forms containing consonant sequences that would violate this constraint. Itô and

Mester argue that, by maintaining these borrowings' source forms, these borrowings raised the constraints regarding faithfulness to consonant voicing, with this raised faithfulness applying to future borrowings as well. In a later epoch leading to many borrowings from Western languages, loanwords carrying phonemic /p/ entered the language and led to a similar raising of the relevant faithfulness constraints and, resultantly, faithful replications of /p/ in those loanwords and allowance for /p/ elsewhere in the Japanese grammar. However, this did not apply retroactively to older Sino-Japanese loanwords, even though they also contained /p/, because they had entered the language and become established before this second contact-induced change. Synchronic analysis of loanwords from different diachronic epochs therefore illuminated how the phonology of Japanese changed as a result of contact and borrowing.

Aside from loanwords retaining sounds or features that are still otherwise not complicit with the borrowing language's phonology, as applied to the native/'core' or oldest members of the lexicon, they also sometimes exhibit unexpected phonological *processes* (e.g., Steinbergs 1985; Shinohara 1997; Kenstowicz 2005; Peperkamp et al. 2008; Broselow 2009). For example, Smith (2006; 2009b) documents that Japanese loanwords in which the source forms contain consonant clusters are adapted to a licit phonotactic structure via epenthesis of a vowel between the originally adjacent consonants; this is unexpected because native Japanese words exhibit deletion to resolve consonant clusters that come about by morphological concatenation. This observation is in line with Paradis and LaCharité's (1997) Preservation Principle contending that loanword adaptation prefers inserting new material rather than deleting or changing phonological features of the source form.

However, research suggests that this preservation can also extend to phonetic-level detail. As another example of a loanword-specific process, Kang (2003) demonstrates that English-to-Korean loanwords often, though variably, exhibit vowel epenthesis after word-final stops (e.g., English *gag* [gæg] → [kæki]), even though Korean phonology does not bar native words from being stop-final (e.g., native Korean 'guest' [kæk]). Kang demonstrates that the



variable pattern of word-final epenthesis in loanwords reflects the influence of a fine-grained phonetic property: The probability of this epenthesis is influenced by the probability of variable word-final stop releasing in English. Kang concludes that the epenthesis is a result of adaptation preserving the phonetic-level cue of stop releasing, since the phonetic system of native Korean leaves word-final stops unreleased.

These observations suggest that loanwords constitute a category that can be meaningful to the sound system: that loanword status, in and of itself, can suppress constraints that native words *are* subject to, or it can sometimes induce processes that native words are *not* subject to. Furthermore, given that loanword status is a result of the context surrounding a word's entry into the language, paying attention to that context is also important to understanding how this feature may be an influential part of the linguistic system. Referring again to the example of how borrowing led to the /f/-/v/ distinction in Modern English, we might consider that French loanwords borrowed into Old English during the Norman Conquest may have been indexed as foreign and therefore phonologically peripheral, allowing for the two sounds to appear in environments where the Old English phonology would have otherwise neutralized the distinction. And this faithfulness, we might imagine, was also enhanced by the association of French at that time with nobility and prestige (Baugh and Cable 2013). Then, as such words became more 'core' to the language and indexed as not foreign at all, so too did this sound distinction they carried with them. A converse approach can also be useful to us: Instead of considering how our understanding of loanword status and adaptation can be informed by the surrounding context, we can examine loanword status as it interacts with language variation and change to inform us about the surrounding context.

### **1.3 Loanword variation**

#### **1.3.1 A not-quite-arbitrary variable**

The field of linguistics largely considers the form+meaning pairings of language to be arbitrary (de Saussure 1916; Hockett 1960). Most word forms are not inherently tied to the

meanings they are paired with, certain pockets of the lexicon (like onomatopoeia) aside (cf. Bergen 2004; Perniss et al. 2010; Dingemanse et al. 2015; Monaghan et al. 2016). And, different languages can be distinguished by what form+meaning pairings they use, alongside grammatical differences. Recursively, we consider differences between language *varieties* to be largely arbitrary as well, where the sociolinguistic variable is simply the use of different surface forms to express the same meaning (Labov 1972, 271). Because of this arbitrary nature, we would say that the difference in form associated with a particular meaning across different groups or contexts can be socially significant, but we wouldn't say that one language or language variety's form is less faithful to the source, because there is no "true"/"original" form or source associated with a particular meaning that one language (variety)'s form can be judged as more similar to.

We can, of course, compare the surface forms of different languages (or language varieties), with such comparisons and their interpretations being foundational to many kinds of scientific linguistic inquiry. We can examine how the shared features of different languages can indicate their shared inheritance and ancestry (e.g., Cavalli-Sforza et al. 1988) or their historical contact beyond a shared ancestry or in spite of a lack thereof (e.g., Gray et al. 2010). Or, as is the cornerstone of sociolinguistics, we can observe how speakers of the same language may still differentiate from or assimilate to others along subtle linguistic variables depending on their contact and/or social alignment (Labov 1972; Bell 1984; Le Page and Tabouret-Keller 1985; Giles et al. 1991; Milroy and Milroy 1992; Eckert 2004). For example, many variable features of a language could be subject to change at any given time, but it is arbitrary and hard to predict what features will actually end up changing, which Weinreich, Labov, and Herzog term the 'actuation problem' (1968, 102): e.g., a vowel shift, a sound merger, a morphological paradigm leveling. Once a change or set of changes has occurred or at least started, two different language varieties may end up resembling each other more than a third, reflecting that those two social groups may have a stronger social connection. But, the source of any change determining the resulting language varieties' forms, was still

arbitrary.

With loanwords, however, a defining aspect is that the borrowing language is also borrowing the sound form of the new word it is incorporating into its lexicon. The form itself is therefore grounded to the form used in the source language when the word was borrowed, rather than being an arbitrary form+meaning pairing or an arbitrarily actuated change. So, when we look across different borrowings of the same word, we *can*, in a sense, say that one adaptation may be closer to the source form than another.

In many cases, such a difference can be attributed to the borrowing language's sound system: e.g., Language A's adaptation of a word borrowed from language X is more similar to the word's form in language X than language B's adaptation of the same word borrowed from language X; this is because language A shares more phonological features and/or phonetic characteristics with language X than language B does. For example, the English word *site* [saɪt] borrowed into French and Italian (in reference to the meaning 'web site') is realized as [sit] and [sito], respectively. The French adaptation is more similar to the English pronunciation than the Italian adaptation: both adaptations exhibit a change of [aɪ]→[i], but the Italian adaptation also exhibits the insertion of a vowel at the end since Italian phonology does not allow a word to end with [t]. This is a recognizably broad, simplified characterization of much of the phonological analysis of loanword adaptation discussed in more detail above (§1.2); but, what this means is that differences in the degree of source form replication might be attributable to aspects of the linguistic sound system rather than some difference in social alignment.

This approach can also sometimes explain differing loanword adaptations between varieties of the *same* language. As mentioned above (§1.2.2), Lombardi (2003) discusses how crucial differences in the phonological systems of different languages can explain their differential substitutions of English /θ/, where some languages adapt this to /s/ while others adapt to /t/. Brannen (2002), however, observes the same differential substitution between European and Quebec varieties of French, which adapt /θ/ in English loanwords to /s/ and

/t/, respectively, in spite of them both sharing the same relevant phonemic inventories, features, and phonological systems. Brannen argues that differences between the two varieties' phonetic realizations of their respective /s/ categories explain why each system chooses a different sound as its optimal replication of the source form. In other words, both varieties are using the most source-like pronunciation they can, but their respective phonetic systems lead to different results when evaluating which sound is more source-like.

However, this approach cannot always fully explain the co-existence of different loanword adaptations within the same language. Boberg (1997; 1999) examines how orthographic <a> in loanwords in English, termed 'foreign (a)', varies in its sound correspondence between /ɑ/ and /æ/ sound categories: e.g., the first vowel of *pasta* or the second vowel of *Iraq*. Regarding British English, Boberg suggests that there are some phonological limitations on which variant is possible in certain environments. Regarding American English, though, Boberg suggests that either vowel category is more readily available, rather than the choice being phonologically dictated; these options are therefore potentially subject to sociolinguistic variation and change. (And, indeed, such variation and change are apparent, as further discussed below.) Like Boberg discusses regarding American English, when one single language or language variety exhibits variation between two different adaptations, in *spite* of both forms being available with respect to the sound system, it seems likely that factors of the social context may be mitigating this more so than factors of the sound system. And, if there is a difference between the two adaptations regarding how strong a replication of the source form each achieves, this may reflect something about the relation between the respective speakers and the source.

It is this kind of case that the current dissertation is largely concerned with. For example, there is notable variation within American English of *Iraq* between [ɪˈɹɑk] and [aɪˈɹæk] pronunciations (Hall-Lew et al. 2010; Silva et al. 2011). The Arabic source form of *Iraq* [ʕiˈraq] is more closely replicated in the [ɪˈɹɑk] pronunciation variant than in the [aɪˈɹæk] variant, where the prior maintains the [ɑ] form of the second vowel and the [ɪ] pronunciation

of the first vowel is arguably more source-like than the [aɪ] diphthong. Since both of these pronunciations are still available in the sound system of American English, we might ask what this variation reflects about the social relation between the source and users of one variant or the other, as well as how this variation comes about. This notion of replication of the source form as a signifier of the relation to or attitude toward the (socially associated) source is previously suggested by work on both loanwords (Weinreich 1968; Hill 1995; Lev-Ari and Peperkamp 2014) and the pronunciation of personal names (Lipski 1976; Kohli and Solórzano 2012; Bucholtz 2016), a similarly not-quite-arbitrary variable. For example, Lipski (1976) discusses how the less source-like pronunciation of *Italian* as “eye”*talian* (an initial diphthongal [aɪ] as opposed to an initial vowel of [ɪ], which would also be available in the English phonology) tends to be observed among “those with a low regard for Italians” (113), the social group to which this word can refer. And Kohli et al. (2012) discuss how the personal names of members of ethnic minority groups in the US are often mispronounced or hyper-Anglicized (e.g., one who pronounces her name *Andrea* as [an'dɪæə] hearing her name pronounced by another as [ˈændɪə]), interpreted as a reflection of how minority groups are ‘othered’.

### 1.3.2 Relevant social factors

Aside from the wealth of studies examining the adaptation of loanwords as a way to understand the inner workings of the linguistic sound system, some attention has also been paid to the variation of loanwords as a way to shed light on the social context and its connection to the linguistic system. Much of this work has explicitly analyzed this variation as variation in the degree of adaptation, with certain groups or situations considered to be using more or less source-like pronunciations (e.g., Poplack and Sankoff 1984; Poplack et al. 1988; van Oostendorp 1997; Thomason 2001; Silva et al. 2011). In some work, this variation of loanwords is referred to as that of ‘(un)nativization’ or ‘integration’. To retain transparency and avoid a potential semantic conflation of the loanword form’s similarity to the source

form with the concept of how frequent or old a loanword's usage may be in the borrowing language, the terminology of 'more source-like' and 'less source-like' will be preferred in this writing. It should also be noted that this terminology refers to the resemblance of the adapted form/variant to the source form; this does not necessarily entail that the speaker, as an individual, cognitively knows of or aims for such resemblance.

There are studies that intuitively show how loanwords can vary between more or less source-like pronunciations due to the intensity of contact between the borrowing language and the source language: e.g., degree of multilingualism at the community level (e.g., Poplack and Sankoff 1984; Poplack et al. 1988; San Giacomo and Peperkamp 2008; Friesner 2009) and familiarity with the source language at the individual level (e.g., Poplack et al. 1988; Kang 2010; de Jong and Cho 2012). There are also studies that observe this variation across factors that could still be considered related to language contact but somewhat removed from the intensity of that contact itself, such as source-directed attitude (Weinreich 1968, 27; Lev-Ari and Peperkamp 2014) and purist language ideology (Poplack et al. 1988; Thomason 2001, 236). For example, Poplack et al. (1988) observe variation in how English-like a pronunciation speakers use for English loanwords in their French speech stream. Individuals who are more proficient in English use more English-like pronunciations, especially for loanwords that are newer and less frequent. They also find that loanwords (especially newer ones) are used more by communities with higher degrees of French-English bilingualism, with this community-level contact intensity being a stronger predictor than bilingualism at the individual level. However, they also observe a difference between two communities of similar contact intensity, suggesting that this may result from these communities' differing attitudes toward anglicisms and regarding the purity of French.

Additional studies suggest loanword variation to pattern with social factors that could be considered even further removed from language contact. The use of more source-like variants appears to correlate with more formal registers, such as with English-to-Dutch loanwords (van Oostendorp 1997) and Arabic-to-Turkish loanwords (Thomason 2001, 73). The

aforementioned variation of ‘foreign (a)’ in American English between [ɑ] and [æ] variants, observed by Boberg (1997; 1999), suggests both an indexation of prestige and a change in progress. The use of [ɑ] appears to carry more ‘linguistic security’ (meaning speakers consider it to be the “standard” or “correct” variant [Labov 1966]), and it seems to be gaining preference as the adaptation of new, incoming loanwords.

Hall-Lew et al. (2010; 2012) look at loanword variation specifically in the second vowel of *Iraq* and find that it patterns with political identity and stance-taking amongst US politicians, with the more source-like [ɑ] variant patterning with liberal identity and stance-taking. Silva et al. (2011), in a follow-up study, do not observe the same political pattern of the ‘foreign (a)’ in *Iraq* and *Iran* amongst non-politician speakers of American English; however, they do observe that those with military experience more commonly use the [æ] variant and those self-reporting as multilingual more commonly use the [ɑ] variant. In their interpretation, they suggest that this may result from “an ‘us-not-them’ mentality” (188) amongst the military community. This resembles an ideology of nationalism. To enlighten our understanding of American nationalism, Banikowski and DiMaggio (2016) sociologically examine the ideologies held by US residents regarding what it means to be American and what pride being American might hold. They identify ‘ardent nationalists’ as those who are more selective about how one qualifies as American, such as having to have been born in the US or having to be White, Anglo-Saxon, and/or Christian. Ardent nationalists also have more hubris (i.e., pride combined with a sense of superiority) regarding the US, such as being proud of US achievements and society or agreeing with the statement that “Others should be like us.” It is also apparent that ardent nationalists hold most strongly to the opinion that one must speak English to qualify as American, and they are highly likely to identify strongly as Republican while unlikely to identify as Democrat.

These findings suggest that nationalist ideology can be relevant to both language and political identity, as well as their connection. Like Banikowski and DiMaggio (2016) observe, this variety of nationalism appears to coincide with political identity, with research

in political and social psychology finding Republicans to hold more strongly to nationalist ideologies and biases (Pratto et al. 1994; Sidanius and Pratto 1999; Jost et al. 2008). Schmidt (1998) discusses how the US ideology about what constitutes national identity favors an ‘assimilationist’ (rather than ‘pluralist’) approach, with this assimilationism being reflected in the dominance of English and the pressure for speakers to linguistically conform. Further research has corroborated a political divide regarding language policy, with Republicans in the US being those who are more likely to oppose bilingualism and support English-only or English-first policies (Citrin et al. 1990; Tatalovich 1995; Barker et al. 2001; Rothstein 2001; Baran 2017). The connection of this nationalist ideology to language may therefore be further reflected in the variation of loanwords, where those who are less nationalist-aligning may be those who are more likely use more source-like (i.e., less assimilated) loanword pronunciations. And, the fact that this intersects with both political identity and language-related ideology may be relevant to the previously observed patterns of loanword variation.

This is where this dissertation situates itself. There are many social factors that may condition loanword variation. Some seem less directly connected to the intensity of contact with the source; however, they are not necessarily orthogonal to language contact or to each other, as Chapter 2 will address in more detail. Furthermore, as discussed above (§1.3.1), this variation regards the replication of the source form and is therefore less of a traditionally arbitrary sociolinguistic variable due to the triangulated relationship of each variant with the source form. This may, by extension, reveal a triangulation between each variant’s respective speakers, groups, or contexts and their social relations to the source of the loanword (or, more broadly, the sources of many/any loanwords).

The first part of this dissertation (Chapter 2) will examine the variation of loanwords in American English more thoroughly. The variable itself will be discussed and reconsidered: The variation of loanwords as a broader group (as previously motivated: §1.2.3) will be examined across diverse source languages and sound variables. Special attention will be paid to the different social factors considered in this section, especially teasing apart US political



identity from them. It is hypothesized that factors like source-directed attitude, broader language contact ideology, and globalist/nationalist alignment may be stronger predictors of loanword variation than political identity, with the variation along political identity being observable as a second-order result (see Silverstein [2003]: ‘indexical order’) of the correlation between political identity and these more direct influences.

However, even if a factor like political identity is shown to be a second-order indexation of loanword variation, it may very well still be meaningfully pertinent to this variation. In an additional study (Chapter 3), analysis will be extended beyond that of how this variation in speech *patterns* with certain social factors; it will further test how loanword variation is *perceived* by listeners who hear it. Attention is paid to how these two phenomena compare, such as whether the same social factors observed to best predict loanword variation are also those social indexations most strongly activated for listeners who hear it.

As Silverstein (2003) and Eckert (2008) discuss, the indexicality of sociolinguistic variation can be complex, fluid, and subject to reconstrual. Social Factor 1 may be a primary influence and, because of this, a stronger predictor of the variation at hand. But, Social Factor 2, even if second-order to Factor 1 as a predictor of variation, may still carry strong associations with the variation of interest and become a meaningful indexation in its own right. In the case at hand, a factor like globalist/nationalist ideological alignment may be a better predictor of loanword variation than political identity. For example, those aligning with a more nationalist ideology may be more likely to use less source-like loanword pronunciations and, while those aligning more with a Republican political identity pattern the same way, much or all of this variation may be accounted for by globalist/nationalist alignment when it is tested as a predictor alongside political identity. But, it may go too far to conclude that the political indexation of loanword variation is a mere byproduct. For one, it may still significantly explain some of the variation observed, if not as much as that explained by considering globalist/nationalist alignment. But, people may still come to strongly *associate* loanword variation with political identity, even if loanword variation

only patterns with political identity in a second-order fashion or as a byproduct of another conditioning factor. This may be especially the case when considering how political identity has recently been increasing in its public salience and factioning in US public discourse (Abramowitz 2013; Westfall et al. 2015).

Accordingly, the study in Chapter 3 will examine listeners' subjective evaluations of loanword variation, both implicitly through a 'matched-guise' design (Lambert et al. 1960; Zahn and Hopper 1985; Purnell et al. 1999; Campbell-Kibler 2007; Yuasa 2010) and by explicitly eliciting subjective evaluations (Preston 1989; Preston 1999; Dailey-O'Cain 2000; Alfaraz 2002). This examination will consider indexations similar to the social factors tested as predictors of variation. This will provide a sense of how strong certain associations with this variation may be, even if such social factors exhibiting this variation in production only do so in a second-order fashion by way of their correlation with other factors.

### **1.3.3 Variation as a result of socially mediated adaptation**

The above discussion of the sociolinguistic variation of loanwords mostly considers the variation of established loanwords (i.e., those that have already been used for a significant amount of time or with a significant frequency and prevalence). However, it is not a far stretch to imagine that this variation may stem from social influences at the incipient stages of a loanword's entry into the borrowing language and its following dissemination and establishment. The variation of an established loanword between more and less source-like variants may reflect how the social factors across which this variation occurs originally mediated the imitation of that loanword's source form.

Studies of loanword adaptation indeed suggest that it is a developing process rather than some one-time instance that cements the resultant form or that repeats anew with every utterance (Haugen 1950; 1953; Kang 2010; de Jong and Cho 2012). For example, Davidson (2007) finds that when native English speakers utter nonce words with illicit consonant clusters (\* /CC /), they produce an excrescent schwa in between the consonants as a result of

trying to produce both but mis-timing their coordination ([C<sup>o</sup>C]). Then in a follow-up study, new participants listen to such utterances and transcribe them, showing a preference to posit a phonological vowel between the consonants (/CəC/). This is in spite of Davidson showing that the excrescent vocoid is significantly different in terms of duration and acoustic quality from phonological schwa. This suggests that 1) loanword adaptation may not be immediate but, instead, a chain-like process, and 2) phonetic-level detail can influence this adaptation as it converges upon a phonologized representation.

We can imagine that this process of loanword adaptation is mediated by social factors. Studies outside the realm of loanword adaptation show that instantaneous, online linguistic imitation is mediated by social factors such as attitudes, shared identities, and power/role dynamics (Pardo 2006; Babel 2009; 2010; Yu et al. 2013; Weatherholtz et al. 2014). And this can influence whether a person or group adopts the linguistic features of another (Labov 1972; Milroy and Milroy 1992; Eckert 2004). There is experimental evidence of a similar phenomenon in loanword adaptation. Lev-Ari et al. (2014) find that when French speakers are exposed to [ʤen:a], a nonce word framed as being Italian, they are less likely to adapt the [ʤ] exposure form to [ʒ] (the usual Italian-to-French adaptation) when told the word is for a kind of ice cream—an area in which Italian products carry prestige—as opposed to a kind of beer—an area in which Italian products are stigmatized. Therefore, imitation of the source form in this simulated loanword adaptation seems to be mediated by the social factor of attitude regarding the source (in this case, domain-specific attitude).

In both of these cases summarized above (Davidson 2007; Lev-Ari and Peperkamp 2014), we might still predict the final outcome on the basis of the sound system alone. Since /ʤ/ is still not considered a member of the phonemic inventory of French and the consonant clusters Davidson examined were illicit in English, we can imagine that it would just be a matter of time (and possibly also the span and frequency of word usage) before the final, regularized adaptations are /ʒ/ and /CəC/ forms respectively. However, it is possible that the social context could slow this process down for, say, a particular group

that maintains stronger contact with the source language or more positive attitudes toward the social groups they associate with it. Or, this could mediate whether loanwords gain a special status such that they are considered exceptional to phonological constraints (as discussed above: §1.2.3), which we could imagine patterning differently across certain groups or contexts. This is similar to what Poplack et al. (1988) observe regarding the pronunciation of English loanwords by French speakers: More established loanwords are more likely to be fully adapted, but certain speakers and social groups may maintain a less adapted, more source-like form longer than others.

This is also relevant to the variation of interest here: multiple adapted forms that are *each* licit in the borrowing language but where one might be considered a stronger replication of the source form (e.g., the variation of *Iraq* between [i'ɪk] and [a'iæk] pronunciations). In cases like this, not only might social factors mediate how far along a certain speaker or group is in adapting the loanword away from the source form; the same social factors might also explain who has settled on an adaptation that is less similar to the source form, in spite of a more source-like form also being available. This may arise from such speakers or groups exhibiting a weaker imitation of the source form along the path of adaptation and, in some cases, that noise in the process leading to a different end result.

A further study within this dissertation (Chapter 4) is intended to address this possibility of loanword variation resulting from socially mediated adaptation. This study experimentally examines simulated loanword adaptation to test if the same social factors exhibiting variation of established loanwords mediate the phonetic imitation of nonce words. Different from the studies above, however, this study novelly divorces loanword status from phonological non-nativeness. Using a phonetic imitation paradigm (Pardo 2006; Babel 2009; 2010; Yu et al. 2013), nonce words are manipulated along phonetic variables in which both ends of the continuum would be licit in American English. Loanword status is achieved by manipulating the surrounding context in which participants are first exposed to the “source form”. The hypothesized effect is that this will pattern in parallel with the variation of es-

tablished loanwords: Speakers more likely to use variants of established loanwords that are more source-like will also be those who more strongly imitate the form of nonce loanwords they are newly exposed to. This is analogous to an apparent-time approach to the study of language variation and change (Labov 1963; Bailey 2002), where language differences between two generations of speakers can be examined synchronically to allow for the inference of a diachronic language change. An observation that variation and adaptation pattern in parallel would lend apparent-time support to the hypothesis that the sociolinguistic variation of established loanwords is a result of the same social factors' influence on the earlier adaptation of new, incoming loanwords.

#### 1.4 Roadmap

The first study in this dissertation (Chapter 2) sets the stage by examining variation in the pronunciation of established loanwords by American English speakers. First, a diverse while interconnected suite of social factors (already foreshadowed above) is identified as potential predictors of this variation. Hall-Lew et al.'s (2010) previous study of the variation of *Iraq* in American English between more source-like [r'ɪɑk] and less source-like [ɑr'ɪæk] pronunciations is the backdrop against which this study is set. Hall-Lew et al. observe this variation to pattern with US politicians' political identity. This observation is taken seriously and dissected rigorously, considering how such a pattern may fall out from and therefore reflect the correlation of political identity with other social factors like source-directed attitude, language contact ideology, and globalist/nationalist ideology. Multiple scholarly studies regarding language contact, loanword variation, and loanword adaptation are considered in the motivation of recognizing these potential predictors, as well as qualitative analysis of public discourse regarding loanword variation.

In a series of experiments eliciting speakers' utterances of variable established loanwords, these social factors are compared alongside that of political identity as predictors of loanword variation. It is hypothesized that similar variation along political identity will be

observed, as well as extending to other loanwords of different sources, diverse sound variables, and less political charge. But, it is further hypothesized that loanword variation will be better explained by these other social factors of source-directed attitude, language contact ideology, and globalist/nationalist ideology when they are considered alongside political identity as predictors. Such an observation will provide support for considering the variation of loanwords with political identity a second-order (Silverstein 2003) reflection of political identity's association with these other more direct predictors.

The next study (Chapter 3) examines how people perceive and think about this variation, as well as how such perception lines up with the way loanword variation patterns in actual usage. It is hypothesized that sociolinguistic perception will parallel how sociolinguistic variation patterns in speech production: e.g., that the use of more source-like pronunciations will be considered as more globalist-aligning as well as politically liberal. However, two methods will be used to examine this perceptual indexation: an implicit matched-guise method and a metalinguistic questionnaire asking explicitly for one's opinions about what loanword variation means. It is hypothesized that perception may similarly reflect a layered nature of this indexicality, predicting that loanword variation may be associated with political identity more strongly when thought about explicitly than any such indexation may be activated during online processing. How listeners *evaluate* loanword variation will also be examined, hypothesizing that the use of more source-like pronunciations may carry prestige and linguistic security ( as Boberg [1997; 1999] observes). The combined considerations of globalist indexation and prestige will advance our understanding of how English and the use/pronunciation of loanwords within it relate to a speaker's capital in what we might consider the global linguistic market (Bourdieu 1977; Piller 2001; Zhang 2005; Blommaert 2010).

The final study (Chapter 4) cycles back to the source, examining how this sociolinguistic variation of loanwords may arise in the first place. The same pool of participants from Chapter 2, who were examined for their pronunciations of established loanwords, are

presented with nonce words framed as loanwords and subsequently prompted to utter them after having heard them. These subsequent utterances are examined for how closely they resemble the exposure form participants first heard uttered. It is hypothesized that this, too, will pattern similarly with the variation of established loanwords: that the social factors conditioning the variation of established loanwords also condition how closely participants replicate the exposure forms of new loanwords they hear. This will therefore shine light on the connection between loanword variation and adaptation, with such a finding suggesting that loanword variation may come about by the initial stages of loanword adaptation being socially mediated by similar social factors.

## Chapter 2

### Variation of established loanwords

#### 2.1 Introduction

While English at present is a language in strong contact with many other languages throughout the world, this contact tends to be fairly light on the English end. Nonetheless, even light language contact can lead to lexical borrowing (Haugen 1950; Thomason and Kaufman 1988; Thomason 2001; Sankoff 2002), and it has done so in the English language. The sociolinguistic variation of these loanwords borrowed into the English language can therefore provide a looking glass into the social dynamics of this contact. And, given that this contact involves a diverse suite of languages, the variation of loanwords from multiple sources can be examined and tested for co-variation as a broad class. This allows us to test whether there are social dynamics that play a role *across* contact languages, rather than considering each loanword (or loanword source) a separate, insulated contact situation.

A loanword, by definition, is the imitation of the sound form of a word from another language (as opposed to a semantic borrowing or calque). Of course, it is rare that this imitation is a completely faithful replica of the source form. Given that the borrowing language's sound system likely differs in some or many ways from that of the source language, the sound form is subject to adaptation which can often be influenced by the borrowing language's phonotactics, sound inventory, or even phonetics (as surveyed in Chapter 1, §1.2.2). This can result in variation across borrowing languages in how the same loanword from a single source language is adapted. However, sometimes there is variation of the adapted forms of loanwords even within a language. Some of this can also be explained by the borrowing sound system, such as by differences between the sound systems of language *varieties* within a language (e.g., Boberg 1997; Brannen 2002). But, much of this variation also seems to be



explained by factors external to the sound system. One intuitively relevant external factor is the degree of bilingualism with the source language, both at the level of the community (e.g., Poplack and Sankoff 1984; San Giacomo and Peperkamp 2008; Friesner 2009) and the individual (e.g., Kang 2010; de Jong and Cho 2012). However, social factors that may be considered even more external to the speaker’s grammar have also been shown to play a role.

Loanwords can vary between pronunciations that are more or less faithful to their source forms even when the sound system of the borrowing language (variety) or the intensity of contact with the source language do not fully account for it. One social factor that may play a role in the variation of loanwords in American English appears to be political identity. Hall-Lew et al. (2010) find that the pronunciation of *Iraq* varies amongst US politicians: Democrat-identifying politicians use the more source-like [ɑ] pronunciation of the second vowel more than Republican-identifying ones, who more frequently use the less source-like [æ] pronunciation. The current study will build on this research, further examining how political identity may condition loanword variation between more and less source-like variants in American English, as well as considering additional social factors that may also condition such variation. Previous research has suggested that attitude toward the language, group, or place of origin can mediate the degree to which loanwords are adapted (Weinreich 1968, 27; Lev-Ari and Peperkamp 2014). This parallels the findings of research not specific to loanwords, where speaking more or less like someone else seems to pattern with the attitude toward or shared identity with that person or associated group(s) (Bell 1984; Giles et al. 1991; Eckert 2004; Babel 2009; 2010; i.a.). Attention will also be paid to how loanword adaptation can be mediated by ideological receptiveness regarding language contact (Poplack et al. 1988; Thomason 2001), as well as considering the influence of an even broader ideological alignment of globalism vs. nationalism—i.e., one’s receptiveness regarding people and/or cultures (beyond just languages or words) they might consider foreign.

These considerations generate multiple hypotheses, which the current study is designed to speak toward:

1. **Political variation amongst non-politicians:** Speakers of American English who are *not* politicians will also exhibit this sociolinguistic variation along the factor of political identity, given that political identity is also (increasingly) salient and factioning in the general public as well (Abramowitz 2013).
2. **A broader pattern than the variation of *Iraq*:** This variation between more and less source-like variants will extend to other loanwords of less political charge and of different sound variables: i.e., This pattern is part of the variation of loanwords as a broad class, rather than a pattern particular to one specific item.
3. **Political variation as second-order:** Source-directed attitude, ideological receptiveness regarding language contact, and one's alignment with a globalist or nationalist ideology are arguably more direct influences on the adaptation of loanwords. It is therefore hypothesized that these may be better predictors of loanword variation when compared as predictors alongside that of political identity.

The following sections will motivate these theoretical considerations and the predictions they generate (§2.2), then turn to empirically testing those predictions (§2.3).

## 2.2 Potential roots to political loanword variation

Hall-Lew et al. (2010) examine the speech of politicians in the US House of Representatives during a 2007 debate regarding whether the number of US troops present in Iraq should increase or not. They find that political identity is a factor exhibiting significant sociolinguistic variation. Republicans are more likely to pronounce the second vowel of *Iraq*, which is a variable 'foreign (a)' (Boberg 1997; 1999; 2014), with the less source-like [æ] vowel while Democrats favor using the more source-like [ɑ] vowel. Hall-Lew et al.'s analysis also considers other factors that could account for some of this variation, such as speakers' US region and variety of American English; however, political identity is observed to be the only significant predictor. They conclude by these findings that political identity is indeed a social factor

that can exhibit sociolinguistic variation, which corroborates and intersects the notions that politics can be relevant to social identity and that language variation reflects the meaningful dynamics and groupings of society.

Furthermore, Hall-Lew et al. (2010) go on to suggest that the phenomenon at hand is not merely a case of sociolinguistic variation occurring across political parties. They state that, for these politicians, “the choice between /ɑ:/ and /æ/ variants is a linguistic resource that aids in the construction of party affiliation and political identity” (98). This statement suggests that the variation of interest may be considered an ‘act of [in this case, political] identity’, as theorized by Le Page and Tabouret-Keller (1985), where “the individual creates for [themselves] the patterns of [their] linguistic behavior so as to resemble those of the groups with which from time to time [they wish] to be identified or so as to be unlike those from whom [they wish] to be distinguished” (181). This idea—that speakers can *use* this language variation with some degree of agency to express and align with certain identities (e.g., Schilling-Estes 1998; 2008; Eckert 2000; Milroy and Gordon 2003)—is somewhat substantiated by follow-up work by Hall-Lew and colleagues (2012), which analyzes intra-speaker variation along this same variable during the aforementioned congressional debate and observes that in some cases this may be attributable to the speaker’s stance-taking in that moment.

The current analysis extends this research program, addressing questions motivated by Hall-Lew et al.’s findings and interpretations. One such extension, as explicitly motivated by Hall-Lew et al. (2010), is the analysis of this variation amongst speakers who are *not* politicians. The meaningfulness of one’s political persuasions to their identity is not particular to politicians whose lives and professions revolve around it; political identity is also potent amongst the general public (Westfall et al. 2015), with seemingly increasing fragmentation (Abramowitz 2013). Therefore, non-politicians are also predicted to exhibit loanword variation along political identity. In addition to broadening the population considered with regard to this variation, a broadening of the variable itself is also considered. Is this particular to

the specific lexeme *Iraq*? Or is it a broader pattern of variation across loanwords of different sources and sound variables between more and less source-like pronunciations? This consideration is theoretically motivated in the following discussion, which provides the primary premise of the current analysis: to consider the additional social factors of source-directed attitude, language contact ideology, and globalist/nationalist ideological alignment. These factors are argued to better explain loanword variation, by nature of being more directly related to language contact, and they motivate the prediction that this variation extends more broadly to loanwords of multiple sources and sound forms. It is hypothesized that these factors are stronger predictors of loanword variation, with such a finding possibly explaining the roots of this variation patterning with political identity. Such a finding might suggest that the political variation of loanwords is a second-order (Silverstein 2003) result and reflection of these social factors, which correlate with political identity but not as one and the same. Furthermore, political identity is examined as a gradient continuum in predicting this variation (especially in Studies 2 and 3: §§2.3.2-2.3.3), rather than a partisan binary. If considering this variation a potential ‘act of [political] identity’, this may predict within-group variation in addition to any variation observable across broadly/binarily defined political groups. That is, those more strongly holding to a certain political identity may more strongly exhibit this pattern if they more strongly wish to resemble a certain political group and indeed consider this linguistic behavior a way of achieving such likeness.

### **2.2.1 Source-directed attitude**

Weinreich (1968, 27) posits that SOURCE-DIRECTED ATTITUDE can influence loanword adaptation. He hypothesizes that, when a more positive attitude is held toward the source language, borrowings from it will likely be less subject to adaptation from their source form. This seems to apply not just to the adaptation of new/incoming loanwords but to the variation of established loanwords as well. For example, Thomason (2001, 73) describes how Arabic loanwords in Turkish vary in pronunciation across different levels of formality, with

more formal styles straying from established adaptations to pronunciations that more closely replicate the Arabic source form—even incorporating Arabic sounds not considered part of the Turkish sound inventory. Thomason interprets this phenomenon as a reflection of the status of Islam in Turkey and the resulting reverence ascribed to Arabic. An additional aspect of Thomason’s observation and interpretation is that this attitude seems associative: One’s attitude toward “the language”, as Weinreich puts it, seems to also pattern with and/or fall out from their attitude toward associated groups, regions, or cultures. (Because of this, the broad term ‘source’ will henceforth be used to collectively refer to this suite and not only to the ‘source language’ specifically). This parallels sociolinguistic research on language change and dialect contact suggesting that attitude toward a variety or associated group(s) will mediate adoption of that variety (Labov 1963; Bourhis and Giles 1977; Giles et al. 1991; Eckert 2004). Lev-Ari et al. (2014) have experimentally demonstrated such attitudinal effects on loanword adaptation. When exposed to [dʒen:a], a nonce word framed as being Italian, French speakers are less likely to adapt [dʒ] to [ʒ] when told the word is for a kind of ice cream—an area in which Italian products carry prestige—as opposed to a kind of beer—an area in which Italian products are stigmatized. In sum, how strongly speakers hold positive attitudes (seemingly either generalized or context/domain-specific) toward the source of a loanword correlates with their inclination to preserve the loanword’s source form.

This hypothesis could be applied to Hall-Lew et al.’s (2010) observed political variation of *Iraq*. First, as in Weinreich’s (1968, 27) hypothesis and many subsequent analyses of loanword variation (Poplack et al. 1988; van Oostendorp 1997; Thomason 2001; Silva et al. 2011), it should be noted that the variable at hand will be discussed as variation in the degree of adaptation: the variation between more source-like and less source-like variants (the terms that will be favored in referring to this variable throughout this study). There is little question that, due to the Arabic vowel inventory, the [ɑ] variant more closely resembles the Arabic source form than the [æ] variant does. Granted, Arabic’s low vowel phoneme is a central /a/, and therefore not a perfect match for either English vowel category; however,

English [ɑ] has been demonstrated to be acoustically more similar to [a] (Fox et al. 1995), so the [ɑ] variant is therefore more source-like than the [æ] variant. (A further motivation for considering the [ɑ] variant of *Iraq*, in particular, as more source-like is the fact that when Arabic /a/ is followed by a uvular consonant (as in *Iraq*) the vowel undergoes assimilatory backing and therefore even more closely resembles English [ɑ].)

We might also consider speakers' *posited* resemblance of a variant to the source form (Smith 2006). When Boberg (1997; 1999) introduces the 'foreign (a)' variable (orthographic <a> varying between [ɑ] and [æ] variants in loanwords), he observes that the [æ] variant carries less 'linguistic security' (Labov 1966): One is likely to use the [æ] variant while still considering the [ɑ] variant to be "correct", more so than the reverse. One interpretation Boberg provides is that speakers may consider [ɑ] a more accurate representation of how the vowel is produced in loanwords' source languages. So, even if speakers are not exactly replicating the source form, they may very well *think* the [ɑ] variant is a closer replication of it.

Regarding Hall-Lew et al.'s (2010) findings, the influence of source-directed attitude could explain the observed variation along political partisanship because this attitude appears to correlate with political identity. Conservatives are found to hold more unfavorable opinions of Arabs and Arab-Americans than those held by liberals, both explicitly (Arab American Institute 2014) and implicitly (Nosek et al. 2007). Conservatives' favoring of the less source-like [æ] variant may therefore be explained as a weaker inclination to preserve the source form of the loan, or maybe even an inclination to diverge from it. Therefore, the differing usage rates of these variants across political identities may fall out from partisans' different attitudes regarding the source of the loanword *Iraq*.

## 2.2.2 Language contact ideology

However, this pattern could also be explained by something broader. Rather than attitude directed toward a particular language, group, place, and/or culture, a speaker's ideol-

ogy about language more broadly could explain their linguistic behavior (Silverstein 1979; Woolard 1992; Jaffe 1999; Kroskrity 2004)—in this case, ideology related to language contact. One holding a less receptive LANGUAGE CONTACT IDEOLOGY may be more resistant to borrowings from other languages, perhaps irrespective of the source and one’s attitude toward it.

For example, Thomason (2001, 236) discusses the contact situation between Montana Salish and American English and how its manifestations are influenced by speakers’ language contact ideology. Thomason describes Montana Salish speakers as highly resistant to lexical borrowing from English: When *semantic* borrowings enter the vocabulary, speakers prefer to use and combine forms native to Montana Salish to represent them, instead of also borrowing the surface form and it therefore being considered a loanword (i.e., a *lexical* borrowing). However, Thomason describes the contact situation between Montana Salish and English as intense, noting how Montana Salish speakers have not been resistant to cultural contact and semantic borrowings. For example, the meaning of ‘automobile’, as Thomason describes (2007, 48), was borrowed into Montana Salish, but given the form *p’ip’úyšn* (native to Montana Salish, meaning ‘it has wrinkled feet’) rather than being an adapted form of English *automobile*, *car*, or the likes. This suggests that it may not be speakers’ attitudes toward the English language or its speakers that constrain borrowing; instead, Thomason discusses how speakers’ ideologies about language contact seem to be at play. This ideology holds a preference to keep Montana Salish distinct from English, which leads speakers to resist manifestations of this language contact that could result in Montana Salish more closely resembling English on the surface.<sup>1</sup> The same could hold for the sound adaptation of loanwords: When lexical (and not just semantic) borrowings *do* enter the language, one with an ideology less receptive to manifestations of language contact may be more strongly inclined to adapt them to forms that feel less foreign.

---

<sup>1</sup>See also Kroskrity (1998) for another example of ‘purist’ language ideology being interpreted to restrict borrowing to only the semantic level and exclude the borrowing of the surface form, and see Poplack et al. (1988) for an interpretation of language ideology explaining different frequencies of lexical borrowing across communities of similar degrees of contact intensity.

This, too, could explain Hall-Lew et al.'s (2010) observed political variation of *Iraq*. Patterns and recent events in US politics suggest a political divide in language contact ideology. Rothstein (2001) reviews the history of policy regarding multilingual education in the US, identifying that Republican legislators are those who tend to withdraw or deny support for multilingual education (see also Citrin et al. 1990; Tatalovich 1995; Barker et al. 2001). Furthermore, this stance does not seem to exclusively regard education but, more broadly, a perceived opposition of language contact to US identity and ideals. This ideology is evident in the following excerpt from an essay about bilingualism in the US (which Rothstein also cites) by Republican Newt Gingrich, published during his tenure as Speaker of the House.

If people had wanted to remain immersed in their old culture, they could have done so without coming to America. [...] Bilingual education slows down and confuses people in their pursuit of new ways of thinking. It fosters the expectation of a duality that is simply not an accurate portrayal of America. [...] Bilingualism keeps people actively tied to their old language and habits and maximizes the cost of the transition to becoming American. [...] The only viable alternative for the American underclass is American civilization. Without English as a common language, there is no such civilization. (Gingrich 1995, 161-162, emphasis added)

Two decades later a similar ideology regarding language contact continues to be espoused by the Republican party, both by new and old faces at the party's forefront. During an interview on CNN, 2008 Republican vice presidential nominee Sarah Palin discusses 2016 Republican presidential primary candidate Jeb Bush's use of Spanish in his campaign. She reiterates the idea that English is the only language that should be spoken in the US.

It's a benefit of Jeb Bush to be able to be so fluent in Spanish, because we have a large and wonderful Hispanic population that, you know, is helping to build America. And that's good. And that's a great relationship and connection that he has with them [...] On the other hand, you know, I think we can send a message and say: You want to be in America? A) You better be here legally, or you're out of here. B) When you're here, let's speak American. I mean. Let's speak English. And that's kind of a unifying aspect of a nation is the language that is understood by all. (Palin 2015, emphasis added, disfluencies removed)

Political events also suggest, as predicted, that this ideological influence extends to loanword pronunciation. While speaking about foreign policy, 2016 Republican presidential



candidate Donald Trump pronounced *Tanzania* as [tæn'zeniə] (McCarthy 2016), while most would consider a pronunciation like [tanzə'niə] or [tænzə'niə] to be closer to how those native to the country would pronounce it.<sup>2</sup> Of interest here is not necessarily what would have been the “correct” way to pronounce these, but the ideologies that this incident and its reception reflect. Trump’s pronunciation sparked some heated dialogue in public media. In her article for *The Guardian*, McCarthy describes academics, left-leaning pundits, and the White House press secretary (under Democratic President Barack Obama) calling Trump’s pronunciation “incorrect” and criticizing him for it. Amidst the media’s scrutiny of Trump, Gingrich again became a representative Republican voice on this issue when he released the following statement defending Trump via Twitter:

Washington elites mock Trump for mispronouncing Tanzania. They don’t get it. He said the most important word correctly: America. He gets it. (Gingrich [@newtgingrich] 2016)

Gingrich’s statement reflects a conservative language contact ideology and highlights that this extends not only to multilingualism but to the pronunciation of foreign words. Gingrich does not argue for the usage or “correctness” of Trump’s pronunciation (even, himself, referring to it as “mispronouncing”); he instead suggests that it is not as important to “correctly” pronounce a placename other than that pertaining to the US.

These examples suggest that broader language contact ideology may be a driving force behind political loanword variation between more and less source-like variants, such as that of *Iraq* observed by Hall-Lew et al. (2010) but also seemingly extending beyond it. The ideology espoused seems to be that one language equals one culture (see Gal and Irvine 1995) and that, for US culture, that language is (American) English (Baran 2017).

---

<sup>2</sup>The dialogue remains unsettled around what *would* have been the most source-like pronunciation of *Tanzania* in both McCarthy’s (2016) article and discussion boards online which include native residents of Tanzania (e.g., Coffeed.com 2009). Personal communication was extended to linguists native to and/or familiar with the linguistic situation of Tanzania, with a consensus emerging that [tanza'niə] is the most common pronunciation to native Tanzanians. Though consultants acknowledge that [tæn'zeniə] is a pronunciation used in the country, this pronunciation may be considered British-English-based and most common amongst non-native Tanzanian residents. Most discussing Trump’s faithfulness to the source form agreed that the [e] vowel (rather than the option between [æ] or [ɑ] alternatives for the first syllable) was the furthest divergence.

Multilingualism within the US, the incorporation of other language forms into English, and attention to preserving the source forms of items that have been incorporated into English are apparent violations of this ideology. It is important to note that Jeb Bush’s use of Spanish in his campaign exemplifies how this factor is not perfectly correlated with political identity; however, his Spanish usage does not appear to have been well-received by the Republican party at large. This suggests that, while political identity and language contact ideology are not one and the same, they do seem to co-vary to an extent that this may explain political loanword variation.

### **2.2.3 Globalist/Nationalist ideological alignment**

Another social influence that may underlie the political variation of loanwords—a broader factor still—is a kind of ideological identity which is distinct from political identity while also seeming to correlate with it. One being more attentive to preserving the source form of a loanword may be a manifestation of an ideology that is more GLOBALIST as opposed to one that is more NATIONALIST. As discussed above (Chapter 1, §1.3.2), US nationalism can be operationalized as entailing exclusivity and hubris, and it is observed to pattern with conservative political identity (Banikowski and DiMaggio 2016). Merry (2016) illustrates how the divide between liberal and conservative US political identities has increasingly come to reflect a contrast between globalist and nationalist ideologies along many facets. For example, in terms of immigration policy, nationalism is less open and more stringent and, in terms of foreign policy, globalism is more strongly favoring of diplomatic approaches. Furthermore, in broader cultural terms, there is a tension regarding how unified/homogeneous the US should be, with a more nationalist ideology holding that cultural uniformity is important to national identity (seemingly parallel to linguistic uniformity, as discussed above: §2.2.2). One example is 2016 Republican presidential candidate Donald Trump at a campaign rally describing the ideal immigrant as one who can “assimilate” (*Los Angeles Times* staff 2016) and other prominent Republican voices like then Kansas Secretary of State Kris Kobach agreeing that

“you have to make [immigrants] become American” and denouncing a “multiculturalism ethic that is [...] repudiating American values” (Kobach 2016). Further research corroborates the correlation of these conflicting ideologies with political identity, such as findings that conservatives have stronger biases disfavoring minorities and those they consider out-group (Pratto et al. 1994; Sidanius and Pratto 1999), stronger biases favoring Americans as a group and the US flag (Jost et al. 2008), and lower scores along the Agreeableness—import assigned to getting along with others—and Openness—interest and comfort in new experiences—social psychological trait spectra (Caprara et al. 2003; 2006; Chirumbolo and Leone 2010; Jost et al. 2003).

It is not a far stretch to analogize this to language: A stronger prescription for immigrants (i.e., foreign people who have entered the country) to culturally assimilate to what may be judged as less foreign might also predict a favoring for loanwords (i.e., foreign words that have entered the language) to assimilate to a form that might feel or sound less foreign. Silva et al.’s (2011) observations hint to this being reflected in the foreign (a) variation of both *Iraq* and *Iran* in a rapid anonymous survey. While they do not find speakers’ political leaning to be a significant predictor, they do find that speakers reporting knowledge of another language are more likely to use the [ɑ] variants while those with military experience are more likely to use the [æ] variants. They suggest that this may reflect a difference in experience; however, their further discussion also considers that it might be a result of the ideological influence of interest here. Those identifying as multilingual might identify as more global while, as Silva et al. themselves suggest, the military community’s preference for the less source-like variants may stem from “an ‘us-not-them’ mentality” (188).

Further evidence suggests that loanword variation both patterns and is perceptually indexed in a way that reflects this ideological variation. Democratic President Barack Obama appears to use resources from other languages to express himself as aligning with a more globalist ideology and persona. In a report for *Politico*, Lee (2009) discusses Obama’s use of more source-like loanword variants, including his pronunciation of *Pakistan* using [ɑ] instead

of [æ]. While Obama’s pronunciation of *Pakistan* is frequently reported on, Lee reports that Obama seems quite consistent in his favoring of more source-like variants of proper names and other loanwords. Obama has also been observed to attempt incorporating at least some formulaic phrase from a non-English language to show respect when he is visiting a non-English-dominant foreign country or when a diplomat from such a foreign country is visiting the US, having done so with many languages including Spanish (Beckwith 2016), French (*Huffington Post* staff 2014), Arabic (*The Economist* staff 2009), Korean (*C-SPAN* 2016), and German (Huetlin 2017). Lee reports that “Pakistanis have told the White House they appreciate [Obama’s more source-like pronunciation of *Pakistan*].” This suggests that Obama is using these features to convey a sense of international diplomacy, and it appears that this is successful.

However, while this may gain Obama social capital in the international, diplomatic linguistic market (see also Chapter 3), it is subject to negative evaluation by some within the US. The following is a post from AboveTopSecret.com (2009), a mainly politics- and government-oriented discussion board site. The post negatively evaluates Obama’s pronunciation of *Pakistan*. It is followed by reply comments, agreeing and adding further interpretation. These interpretations index Obama as self-distancing from the US and aligning with Pakistan and Islam.

**Original post:** “[...] It’s almost like he deliberately pronounces it different than any other American to prove to his Muslim brothers he’s more Muslim hip than the rest of us or something. [...] Burns me up!” –genius/idoit

“YES, YES, YES!!!! I’m glad you brought this up! Drives me nuts to hear how he says it. This is America. We speak English (some of us still do at least.) We don’t say PAH-kee-stahn. We say Pack-ist-an.” –IamLael

“Its either Pack-ist-an or nothing I say. Either you are with us (Pack-ist-an) or you are against us (PAH-kee-stahn) [...]” –Dorian Soran

While Obama may be wanting to align with Pakistan in a diplomatic sense, it is doubtful that, as President of the US, he is intending to distance himself from the country he represents. Rather than an oppositional alignment between these two places, this linguistic choice of Obama’s (and the others mentioned above) may instead reflect an inclusive,

more globalist alignment. The following comments further the discussion in such a direction, adding an interpretation of Obama’s pronunciation as globalist. The first commenter (negatively evaluating the pronunciation) interprets an intent by Obama to be well-perceived on a global scale and considered “cultured” and “international”. The following commenters (positively evaluating it) suggest a similar interpretation. They opine that Obama is pronouncing *Pakistan* the way Pakistanis would like it to be pronounced. They interpret this as being respectful (as opposed to “egotistical”). Therefore, Obama may be more willing to linguistically cater to international audiences and constituencies, in spite of it costing his image amongst others (perhaps especially US nationals).

“i’ll add it to my list of reasons I don’t like him as president. Why Not. He is kissing to much foreign butt anyway and it does sound pretentious. I think our list of real issues is far bigger than this anyway, concern at being perceived as cultured... better feared and tolerated than sucking up and taken advantage of...he’s more concerned about international image than domestic issues [...]” –mopusvindictus

“You mean, the right way? :) Wouldn’t you want people in other nations to pronounce the name of your country the way it’s intended to be? Or should we just make up our own names for countries entirely? Why should he mispronounce the name of the country to make you feel better about yourself?” –ZombieOctopus

“Do you have any idea how the rest of the world cringes when Americans say it? [...] Thank God he’s taken the time to find out how to pronounce it rather than just put his own egotistical spin on it. Its like if my name is Ricardo and someone just blatantly keeps calling me Richard. Its just clueless and arrogant really.” –zazzafrazz

This globalist alignment also appears to be indexed as upper class and prestigious, with the previous commenters also ascribing a sense of correctness to the more source-like variant. The commenters above who positively evaluate Obama’s use of [ɑ] suggest that his pronunciation is “the right way”, and that he “[took] the time to find out how to pronounce it” rather than “mispronounce” it. This suggests a greater sense of linguistic security in the [ɑ] variant, mirroring Boberg’s (1997; 1999) findings discussed above. While the other commenters negatively evaluating it say that it’s un-/anti-American, they do not evaluate it as “wrong” or “incorrect”. They do, however, label it as “pretentious”. So, the use of more source-like variants seems to be associated with prestige, both positively (“correct”)

and negatively (“pretentious”). This calls to mind Preston’s (1999) distinction of whether speakers evaluate a language variety as “correct” or “intelligent”, representing the factor of standard prescription and linguistic security, and whether they evaluate it as “pleasant”. Here, a “pretentious” evaluation suggests the orthogonality of these evaluations, where one may consider a form to be prestigious along a broader sense of the standard (here, ‘prestige’ meaning ‘overt prestige’ as opposed to ‘covert prestige’) while still not being evaluated as pleasant or in-group.

A final comment (of negative evaluation) exemplifies both indexations, framing Obama’s pronunciation as simultaneously global (“exotic”) and prestigious (“ostentatious”, “pretentious”). The commenter also says that what they like about Republican Sarah Palin is the way she uses the [aɪˈræk] pronunciation of *Iraq*, because “When in Rome, do as the Romans.” In other words, her use of the more adapted and less source-like pronunciation is interpreted as more strongly aligning with an American identity and audience.

“President Obama’s overly exotic pronunciation of Pakistan is made even more ostentatious in the same sentence as the rather mundanely pronounced "Afghanistan." Isn't there a more pretentious way to pronounce Afghanistan? How do the Afghans say it? Ahf-GAH-nee-stahn, perhaps? When in Rome, do as the Romans. Here in the USA we say PAK uh-stan. We also say MEK-si-koh, and Frans – and JUR-muh-nee. Maybe we should start calling it Deutschland instead. What I like about Sarah Palin is the way she says "Eye-raq." ” –linux2216

These two indexations of globalism and prestige/class may very well be connected. And, like political indexation, this prestige/class indexation of loanword variation may also be a second-order reflection, inheriting this variation because of its connection with globalism. Someone more global-*ist* might also be more global: That is, they may have more mobility and less local tie. And someone more global generally has to have the economic resources for it, leading globalism to be associated with socioeconomic status (SES). Due to this association of globalism with SES, expressions of globalism may be perceived by fellow nationals as the out-grouping of oneself in a direction of overt prestige—recall the “pretentious” label combined with an un-American indexation in the discussion board comments

above.

Therefore, this SES-related indexation of loanword variation would be a secondary result, stemming from such evaluations of globalism as pretentious. As theorized above, the globalist ideology is arguably a direct influence on loanword variation. It is also arguably a more direct influence than prestige or upper class identity: One's identity as upper class, on its own, does not as easily explain or predict the preference of more source-like pronunciations of loanwords. We can further imagine a speaker who associates with wealth and mobility but not with globalism. If Donald Trump is an example of this, the qualitative analysis of his pronunciations throughout this paper would support the point that wealth and mobility alone might not as strongly predict that one will use more source-like loanword variants, as opposed to their globalist/nationalist alignment.

But, the association of globalism with the upper class may explain why the use of more source-like loanword pronunciations can be indexed as prestigious. This is parallel to how it being indexed with liberal political identity may also stem from globalism, with the approbation of globalism seeming to correlate with liberal identity (as discussed above: §2.2.3). The current analysis focuses more on comparing political identity to the more direct social influences proposed in this paper, given that any indexation or patterning with SES may also be explainable as second-order. SES is therefore not given primary, nuanced attention in this study; however, somewhat relatedly, Study 2 (§2.3.2) will give some attention to how 'mobile' participants identify, and Study 3 (§2.3.3) will consider both participants' mobility and how strongly they express prescriptivist views of language.

To summarize this entire section, there are multiple social factors that seem directly related to language contact and may therefore influence the variation of loanwords in US English between more and less source-like pronunciations. One may be more strongly inclined to replicate the source form of a loanword when they hold a more positive attitude regarding the source language and/or associated groups, places, or cultures. We can also expect that if one holds an ideology regarding language contact that is less receptive to manifestations

of it, they may more strongly prefer to adapt or diverge from more source-like pronunciations of loanwords that have entered the language (regardless of the source). Another possibility is that this is an artifact of (and/or means of expressing) one's alignment with a globalist/nationalist ideology, with those more globalist-aligning more strongly inclined to preserve loanword source forms as a kind of deference to the associated sources. (Or, as discussed further at the end of this chapter, this could also reflect one's inclination to seek the source form and/or their prior opportunity to know the source form.) And, these factors appear to correlate and be co-indexed with US political identity, so they may therefore underlie Hall-Lew et al.'s (2010) previously observed loanword variation along political identity. That is, political identity exhibiting this variation may be second-order (Silverstein 2003), having inherited this variation due to its connection to these more direct social influences. Yet, while all of these factors have been demonstrated to pattern with political identity, they are not necessarily one and the same. It is therefore worth testing how these factors compare with political identity as predictors of loanword variation, with the current analysis testing the hypothesis that these social influences will account for loanword variation better than political identity itself.

## **2.3 Empirical analysis**

### **2.3.1 Study 1: Random sampling questionnaire**

This study is a preliminary examination of whether Hall-Lew et al.'s (2010) observed variation of *Iraq* across political identities extends to non-politicians. Seeing that politicians can be seen as having their own ways of speaking (Safire 2008), and that they appear to exhibit strategic style-shifting to be indexed in different ways depending on the audience and context (Flores-Bayer 2015), it's possible that this variation is limited to the speech of politicians, whose political identities play an integral role in their lives and careers. There is support for this hypothesis, with Silva et al. (2011) finding political identity to not be a significant predictor of *Iraq* and *Iran* variation in a study of non-politicians. On the other hand, we



know that political identity can be meaningful to non-politicians as well (Westfall et al. 2015), such as its increasing prominence and factioning amongst the general American electorate (Abramowitz 2013). Political variation amongst non-politicians may, therefore, still be observable (and perhaps more so, years after Silva et al.’s study). This study also extends the analysis to the variation of other placenames from different sources and of less political charge. The correlation of political identity with broader social influences like language contact ideology and globalist/nationalist ideology might predict the political variation of loanwords to hold across the board. On the other hand, the variation of *Iraq* could be best attributed to source-directed attitude, which also seems to correlate with political identity. If this is the case, we might predict the variation of *Iraq* to be the only placename exhibiting a pattern with political identity, or that exhibiting such a pattern most strongly.

### 2.3.1.1 Methods

A random sampling questionnaire was conducted in the lower Manhattan area of New York City as part of a group project in a Sociolinguistics Field Methods course. People sitting in public parks were approached and asked if they would like to participate. Sampling (N=100) was balanced for age (half of participants 18-30; half over 30) and gender (half female-presenting; half male-presenting). Ethnicity was controlled for, sampling only within the white population so as to increase the likelihood of sampling subjects who identified as Jewish (for other analyses besides that discussed here).

The task was guised to participants as a “geography quiz”. Participants were given a handout with a list of placenames and a prompt asking them to list the places in order by how hot they think each place gets in the summer. (The list and prompt are provided below.) As participants said each placename, the researcher auditorily coded and recorded their pronunciation variant. The three placenames of interest here (with the {more source-like ~ less source-like} variables used for auditory noted) are *Iraq* ([ɑ]~[æ]), *Chile* ([e]~[i]), and *Quebec* ([k]~[kw]). The other placenames included are also variable; however, other

complicating factors of both form and context led to the decision to exclude them from this analysis. (For example, the variation of *Tokyo* between [...jo] and [...io] forms was deemed too unreliably distinguishable when doing rapid auditory coding in an outside environment. The variation of *Beijing* carries a different kind of complication, with the [ʒ] version actually being a ‘hyperforeignism’ [Janda et al. 1994] and it therefore being unclear whether one’s use of [ʒ] is the speaker’s adaptation from a (mis-)posited source form or their preservation (and perhaps flaunted knowledge) of what actually is a closer replication of the source form. And some words, like *Shanghai*, simply showed little to no variation in this dataset.)

Rank these places from hottest to coolest summer.  
(If you don’t know the place, just guess.)

GROUP A	GROUP B
Albania	Nevada
Beijing	Pakistan
Chile	Quebec
Haifa	Shanghai
Iraq	Tokyo
Libya	

Geography Quiz (guised placename elicitation)

Participants were asked to look at each ‘Group’ and say the placenames in the order they decided to be their answer. Pronunciations were auditorily coded and recorded on the spot. All participants did both Groups. (The division was designed to reduce difficulty and processing load.)

After the questionnaire, participants were asked for some demographic information. Of peak interest to this analysis was the question used to elicit political identity: “How would somebody else label your political views? Liberal? Conservative? Other?” Other questions of interest asked about participants’ education level, language experience, travel experience, and US regional identity.

**2.3.1.2 Results**

Figure 1 provides a raw summary of the results: 246 utterances by 88 participants. On the left, these are grouped by political identity as liberal, moderate, or conservative. Most all participants who chose the ‘Other’ option said “Moderate”; otherwise they were excluded.

Some participants were also excluded due to demographic factors like not having grown up in an American-English-dominant language setting or having visited a place of interest for an extensive period of time. On the right, results are grouped by word. As is likely first apparent when viewing Figure 1, recruitment garnered more liberal-identifying participants than conservative-identifying participants. This is not entirely surprising in New York City (*Politico* 2016); however, see §2.3.2.4 for more in-depth discussion of how this might not be solely a reflection of sampling. As is further apparent, liberals use the more source-like pronunciation variants more frequently (51%) than conservatives (14%), moderates patterning with liberals (47%). It also appears that *Quebec* has an especially strong preference across speakers for its less source-like variant ([kwəbək]) compared to the other placenames (63%; cf. *Chile*, 44%; *Iraq*, 49%).

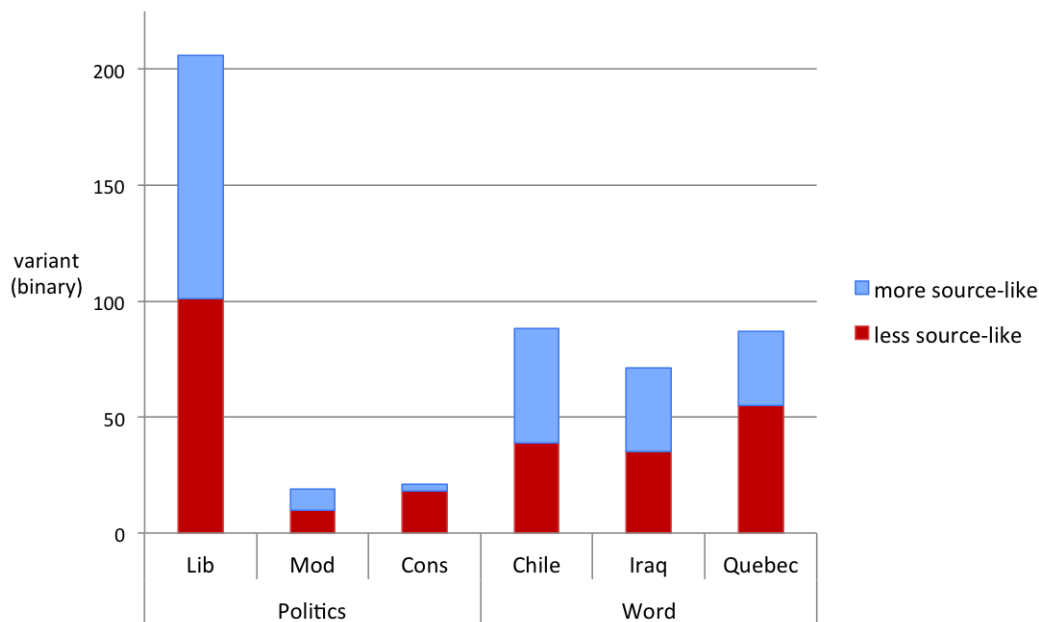


Figure 1: Random sampling questionnaire: raw results

Utterances of interest elicited from the geography quiz. Blue represents the more source-like variants: *Chil[e]*, *Ir[a]q*, [k]e**bec**. Red represents the less source-like variants: *Chil[i]*, *Ir[æ]q*, [kw]e**bec**.

The results were submitted to a step-up/step-down logistic regression analysis in Rbrul (Johnson 2009), testing for significant predictors of whether tokens used the more source-like or less source-like variants. Factors included in the analysis were age, gender, education,

political identity, word, and interviewer. The factors of age, gender, and education were not found to be significant and will not be discussed further. Figure 2 provides a visualization of the Rbrul weights given to the different conditions within the significant factors of political identity and word, where a weight closer to 0 means stronger preference for a more source-like variant and a weight closer to 1 means stronger preference for a less source-like variant.

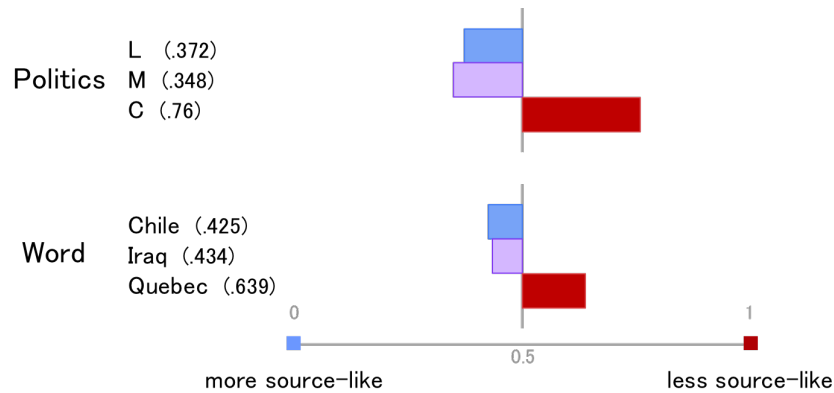


Figure 2: Random sampling questionnaire: Rbrul results

Weights assigned to each factor by the Rbrul model. A weight closer to 0 means a higher likelihood for an utterance of this condition to use the more source-like variant. A weight closer to 1 means a higher likelihood for an utterance of this condition to use the less source-like variant. A weight of 0.5 means that a certain condition is not found to show a preference between either of the variant categories.

The effect of political identity is in the expected direction: Those identifying as liberal show a preference for more source-like variants; those identifying as conservative show a strong preference for less source-like variants. However, those identifying as “moderate/other” do not fall between these two groups but in line with the liberal-identifying. What this may suggest is that these participants indeed have a shared identity and linguistic usage with those identifying as liberals, but political identity itself is not what they align with. (Instead, this could be other ideologies and attitudes like those considered above: §2.2.)

The effect of word suggests that *Chile* and *Iraq* show slight tendencies favoring their more source-like variants, while *Quebec* shows a stronger preference for the less source-like variant. That is, one might use the more source-like variants of *Chile* and *Iraq* while still using the less source-like variant of *Quebec*, likelier than the reverse. It is hard to attribute this effect to contact strength, given the relative prestige and historical prevalence of French

as a contact language with English; as Janda et al. (1994) point out, this actually leads French loans to have more faithful adaptations and for their adaptation patterns to even overgeneralize to loans from other source languages. It is also hard to attribute this to source-directed attitude, since we might therefore predict *Iraq* to show the strongest dispreference of its more source-like variant. This is because, while conservatives are shown to have stronger anti-Arab attitudes (Arab American Institute 2014) and biases (Nosek et al. 2007) than liberals, these are still demonstrated to be held across both groups (Nosek et al. 2007, Cashin 2010). A possible alternative explanation is speakers' perceived degree of foreignness of the word, with *Quebec* pertaining to neighboring (and fellow English-majority) Canada and therefore not being considered as foreign a placename as the others. Or, this might be attributable to the relatively low frequency of *Quebec* and pronunciation defaulting to the usual <qu...> → [kw...] spelling-to-sound mapping.

As discussed above, the primary question with regards to the effect of word is its interaction with the effect of political leaning. A hypothesis that loanword variation stems from source-directed attitude would predict the variation of *Iraq* to show the strongest pattern with political identity. Figure 3 provides proportional cross-tabulations to show how this variation patterned with political identity across the three words. The political effect appears consistent across all three words: In each case, liberals prefer the more source-like variant more strongly than conservatives do. However, it appears that the political effect is actually weakest for the variation of *Iraq*, compared to the variation of the other words. While there is not enough data to allow for a statistical interaction test, these results suggest that the variation of *Iraq* did not carry the weight of the main effect of political identity that was found to be significant above.

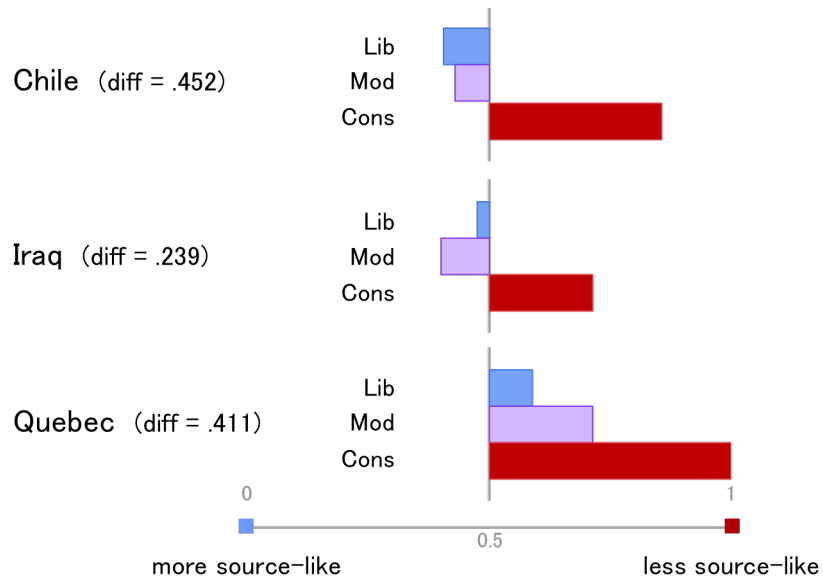


Figure 3: Cross-tabulations: political identity  $\times$  word

Proportional cross-tabulations between political identity (liberal = blue; moderate = purple; conservative = red) and word. A proportion closer to 0 means that a greater proportion of utterances in this condition used the more source-like variant. A proportion closer to 1 means that a greater proportion of utterances in this condition used the less source-like variant. For each word, the difference ('diff') is noted between the liberal and conservative conditions, to demonstrate how strong the political effect was.

While not central to the current analysis, attention is also warranted regarding some additional factors. Recall from above that an effect of interviewer was also found to be significant. Of the five interviewers who conducted this questionnaire, the data from two show an overall skew toward the more source-like variants (Rbrul weights of .338 and .389, respectively), two did not show much of an effect (.493, .552), and one showed a skew toward the less source-like variants (.721). While this could be a random outcome, it could also be meaningful, as interviewer effects have been observed in previous sociolinguistic studies (e.g., Rickford and McNair-Knox 1994; Cukor-Avila and Bailey 2001). This could be the result of a sampling bias, either in terms of the interviewers' selection of potential participants or people's willingness to participate. Regarding interviewers' selection, while each interviewer did balance for age and gender presentation, it's possible that factors like the disposition or presentation of a potential participant may have influenced an interviewer's decision to approach for recruitment, and such factors might not be orthogonal to those of influence considered here (e.g., presenting as more open by their disposition or as more liberal by their

fashion choices). Or, conversely, a recruit's decision to participate may have been influenced by the disposition or presentation of the interviewer who approached them.

It's also possible that such an effect came after recruitment in the form of accommodation (Giles et al. 1991), with a participant perhaps indexing their interviewer with a relevant social attribute and resultantly feeling either less comfortable or especially motivated to use a particular variant. It is notable that, of the two interviewers who showed a skewing toward the more source-like variants, both were females and one was Afro-Latinx while the other was French-accented. It is possible that participants felt a stronger motivation to use more source-like pronunciations due to these indexicalities of their interviewers, while they may still have used less source-like variants with other interviewers. The other three interviewers were white males. It may be notable that the interviewer whose collected data showed a skewing toward less source-like variants had features of the Northern Cities Shift. While interviewers did not say the words of interest, themselves, participants may have picked up on this interviewer's regional dialect. It may be that this was also a case of accommodation, where listeners associated the use of less source-like variants with the Midwest and were more strongly inclined to use such variants upon indexing their interviewer with this dialect region.

Another factor recorded during participation was a participant's US regional identity. Hall-Lew et al. (2010) did not find speakers' US regional identity or regional variety of English to significantly influence the variation of *Iraq*; however, they did consider that its correlation with political identity might explain such a pattern. While the data from this study do not allow for statistical analysis in teasing these factors apart, attention was paid to whether there was a conflation of US regional identity with either political identity or loanword variation. Neither of these possibilities appears to be the case. For example, there were participants who exclusively used the less source-like variants but who identified as liberal and from liberal-trending places like New York City and California, and a majority of conservative-identifying participants did not also identify with conservative-trending US

regions.

In sum, the results of this study suggest that language variation along political lines is not exclusive to politicians: Non-politicians exhibit political language variation, too. Results also suggest that this variation may best be considered a broader variation of loanwords as an entire class between more and less source-like variants, seeing that this effect extends to additional foreign placenames of other sources and less political charge than *Iraq*, as well as other sound variables. The results also provide further support to the hypothesis that this political variation may be falling out from other social factors. For one, it is apparent that those who identify as politically moderate do not fall squarely in between those who identify as liberals or conservatives, suggesting that some other factor(s) may better explain this variation. Furthermore, attention to the the particularities of each word suggest that the broader social factors may best explain this variation. When examining the interaction of word with political identity, the variation of *Iraq* does not appear to carry the weight of the political variation observed, suggesting that source-directed attitude may not be the root of it. (Though this is assuming, while not having directly tested, that Arab/Middle East-directed attitude was that most strongly correlated with this participant pool's political identities, as motivated above: §2.2.1.) And, the main effects of each word suggest that *Quebec* shows an especially strong preference for its less source-like [kw...] variant, with one possible explanation (as discussed above) being that this placename feels less foreign to speakers. This further motivates examining and comparing the above proposed social factors of interest with political identity as predictors of loanword variation, which is done in a more direct method in the next studies.

### **2.3.2 Study 2: Directly measuring and comparing factors of interest**

This subsequent study extends the analysis of loanword variation to further variable place-names still, as well as employing a Likert questionnaire to capture how participants identify along the various social factors of interest. This method allows for a more rigorous and nu-



anced approach to considering these social factors, by allowing for direct comparison of their effects on any observed variation and by treating each factor as a continuum rather than a categorical binning. Finally, this study additionally incorporates an Implicit Association Test (Greenwald et al. 1998) to examine Arab-directed attitude. While the Likert questionnaire explicitly elicits such an attitude measurement, it was hypothesized in the design of this study that an Implicit Association Test (described in more detail below) might more accurately capture participants' attitudes that they may not be as comfortable explicitly espousing in a questionnaire.

### **2.3.2.1 Methods**

For this study, participants were seated in a sound-attenuated area. On the desk in front of them was a computer screen, keyboard, and mouse. Participants' speech was recorded using a Shure SM35 headworn microphone connected to a Zoom H4n digital audio recorder (44.1kHz, 24bit sampling). This was part of a larger study, with a preceding task involving listening to short stories and reading short sequels aloud, which was a pilot version of the simulated loanword adaptation experiment to be discussed in Chapter 4. The task of interest here is a wordlist reading task that participants performed. A word was presented on the screen and then automatically advanced to the next, with the participant reading each word aloud. (Each word displayed for 1.25s with a 0.5s blank screen in between.) The list of words was randomized and cycled through three times, with this randomization new for every cycle and every participant. The variable loanwords of interest are provided in Table 1, along with their variants that participants' pronunciations were subsequently auditorily coded for. There were 47 words in total; the rest were fillers not analyzed in this study (e.g., *deck*, *fatigue*, *gardenia*, *gazette*).

<u>Chile</u>	[e]	~	[i]	<u>Quebec</u>	[k]	~	[kw]
<u>Hawaii</u>	[ʔi]	~	[i]	<u>Samoa</u>	[ɑ]	~	[ə]
<u>Iran</u>	[ɑ]	~	[æ]	<u>Tokyo</u>	[j]	~	[i]
<u>Iraq</u>	[ɑ]	~	[æ]				

Table 1: Variable words of interest: more source-like ~ less source-like

These are the words of interest that were elicited in the word list reading task. The parts underlined are those used in the auditory coding of utterances between more source-like (left) and less source-like (right) pronunciations. Words are spelled as they were orthographically displayed to participants (e.g., *Hawaii* rather than *Hawai'i* was used to avoid a priming effect).

After the word list task was an Implicit Association Test (henceforth, ‘IAT’). The IAT, as developed by Greenwald et al. (1998), is meant to test for and measure the implicit biases one might have in associating certain categories with certain attributes. Participants are directed to associate two categories with two different attributes respectively. This is done by assigning one category + attribute pair the same response key on a keyboard, and the other pair to another key. Participants then perform a sorting task by being randomly exposed to stimuli representing the categories and attributes of interest and pressing the respectively assigned key. Participants’ speed and accuracy in this sorting task is used to measure how much bias they have in associating a certain category with a certain attribute.

For example, Greenwald et al. (1998) tested for participants’ implicit negative biases regarding African Americans. One key was assigned to both the category of African Americans (using African American-indexed first names) and the attribute of pleasant (using word stimuli of positive evaluations: e.g., *happy*, *lucky*), thereby associating the two. For the other key assignment, they would associate the category of White Americans and the attribute of unpleasant. Stimuli would appear randomly on the screen and participants were to sort them by pressing the assigned key as fast as possible. They would also perform the task with the reverse category + attribute associations. Participants were found to respond slower when associating African American sounding names with the pleasant attribute than when associating them with the unpleasant attribute. This suggested that participants were less used to associating African Americans with pleasant attributes or, vice versa, more used to associating African Americans with unpleasant attributes: i.e., they had a negatively bi-

ased association regarding African Americans. Similar results have been replicated regarding multiple other minority groups and using a range of stimulus types, such as images instead of words (Project Implicit 2016). Furthermore, this kind of test has been used in previous sociolinguistic research, examining biases in the perception of linguistic stimuli (Campbell-Kibler 2012; Pantos and Perkins 2013; Callesano and Carter 2016; Laturus 2018) and biases regarding certain social groups/categories to test how they influence language-related behavior, like imitation (Babel 2009; 2010) or perceptual accentedness rating (Yi et al. 2013).

In the study at hand, the IAT component examined participants' biases regarding Iran (meant to access an Arab-indexed categorization) relative to those regarding Samoa (meant to access a foreign-indexed categorization, while possibly more neutral or positively charged). The task tested participants' reaction times in sorting stimuli between two different response keys while using different category + attribute pairings assigned to each key. Faster reaction times suggest stronger implicit association. For example, a participant responding faster when pairing Iran with negative attributes than with positive attributes is interpretable as them having negative-biased associations regarding Iran (and, in this case, likely generalized to the Arab Middle East, Arab and/or Muslim people, etc.). For each category, stimuli consisted of stock images, with respectively paired compositions: e.g., a group of women dancing while wearing traditional garb, a landscape indexable to the respective region. For attributes used to examine associations with these categories, word stimuli were displayed to represent positive attributes (*good, joy, love, pleasant*) and negative attributes (*bad, agony, harm, nasty*). The experiment administered was a tailored version of the Open IAT (Stafford and Scaife 2015) developed for PsychoPy (Peirce 2016) and made available by the Open Science Framework (<https://osf.io>).

Finally, a Likert agreement questionnaire elicited participants' alignment with statements regarding the social factors of interest. In designing this questionnaire, each factor was considered a continuum between two ends: political identity (liberal ~ conservative), place-directed attitude (Middle East favoring ~ Polynesia favoring: generalized from the

more specific Iran ~ Samoa conditions of the IAT), language contact ideology (more receptive ~ less receptive), globalist/nationalist alignment (globalist ~ nationalist), and mobility as it may relate to or be influenced by socioeconomic status (more mobile ~ less mobile). Participants would see a statement and be asked to respond on a 7-point scale where the leftmost option was “strongly disagree” and the rightmost was “strongly agree”. Table 2 provides example statements of each aspect, with pairs in which stronger agreement ratings would represent opposing ends of the continuum; multiple additional statements were used for each (29 in total, provided in Appendix A). Statements were semi-randomized, with attention paid to spacing statements that were similar or opposites. Every participant encountered the same order. A participant’s responses to all statements within each social factor grouping were combined to result in a composite score representing where that participant fell along each continuum.

FACTOR	STATEMENT
<hr/> <hr/> Political identity:	
liberal	My political views are usually liberal.
conservative	I have a lot of Republican friends.
<hr/> Place-directed attitude:	
Middle East favoring	I would like to travel to somewhere in the Middle East.
Polynesia favoring	I don’t consider Polynesia a dangerous region to travel to.
<hr/> Language contact ideology:	
more receptive	I enjoy learning other languages.
less receptive	Immigrants who move to the US should be required to learn English.
<hr/> Globalist/Nationalist alignment:	
globalist	I often fantasize traveling to new places around the world.
nationalist	I don’t keep up to date much regarding global current affairs.
<hr/> Mobility:	
more mobile	I have traveled to a diverse range of places, compared to most.
less mobile	I haven’t had the opportunity to travel much.

Table 2: Example Likert agreement statements

It was hypothesized that the social factors measured in the IAT and Likert Questionnaire tasks might significantly condition participants’ linguistic behavior. Those with a more negative implicit bias regarding the Arab Middle East might be those especially favoring the less source-like pronunciations of Arab-indexed loanwords. This might also be the

case for those who explicitly express such negative associations in the Likert Questionnaire, though participants might not be as willing to express such sentiments explicitly (hence the inclusion of the IAT). Other factors might also condition loanword variation: Those who use less source-like pronunciations might be those who more strongly align with a conservative political identity, a less receptive ideology regarding language contact, a more nationalist ideology, or less mobility.

Recruitment took place in and around New York City, resulting in 27 participants. Criteria for eligibility regarded age and language experience. Eligible participants were 18-35 years of age to control participants as fitting into the ‘Millennial’ generation, especially warranted seeing that the recent decades’ increase in political polarization in the US is evident not only amongst politicians but amongst the general US population and electorate as well (Abramowitz 2013). Participants also had to be native, monolingual speakers of American English who did not have beyond two years of college-level instruction of a non-native language.

Emails were sent to course rosters and participant pools at institutions of post-secondary education in the area, asking for participants in a language study involving “reading out loud, a cognitive task, and a questionnaire”. Explicit mention of politics was avoided so that participants would not enter the experiment with politics already in mind. Such awareness could lead them to bring such an identity more to the forefront than it would be otherwise, possibly affecting their performance. This, however, made it difficult to recruit conservatives, a less populated group in the NYC metropolitan vicinity (*Politico* 2016). After 20 participants completed participation (balanced for gender), preliminary analysis suggested that none of them held conservative political identities, though a full spectrum within the moderate-to-liberal half was represented.

Additional recruitment efforts were made with attention to recruiting more conservative leaning participants. One effort was friend-of-a-friend networking. After completing the experiment, participants were debriefed and asked if they had contacts in the area who

they thought identified as politically conservative. It was requested that they spread the recruitment efforts particularly toward such contacts, though that they crucially not be told they were being recruited because of conservative political leaning. Another effort was on-site recruitment. The study was run at an institution in an area demonstrably more conservative-populated (*Politico* 2016), near but not within NYC proper. At that institution, flyers were given to passers-by which led participants to a neighboring building in which the study was administered. A final effort was a more directed networking attempt. Contact was made with someone holding a leadership role in an organization for young conservatives, with a friend-of-a-friend connection to the author. This contact was asked to participate in the experiment (again, with recruitment never explicitly mentioning politics). After completing the experiment, the participant was debriefed and commissioned to bring more conservative participants to the study who might be known through this organization, with a remuneration being offered for this endeavor.

While these efforts did garner more participants, it only recruited a few who identified as politically conservative. As argued above, this does not necessarily invalidate this dataset's ability to speak to the questions at hand. If we consider loanword variation an act of political identity (as suggested by Hall-Lew et al. 2010), this might predict the effect of political identity to hold regarding the *degree* of such an identity's strength *within* one side of the spectrum: i.e., Those who identify more strongly as liberal might be more likely to use more source-like pronunciations than those who identify as more moderate. However, this is a notable challenge which will receive further discussion at the end of this section (§2.3.2.4).

### **2.3.2.2 Statistical analysis**

Before turning to the results, a precursor explanation of the statistical analysis that will accompany those results is warranted. First, each individual factor of interest will be examined as a lone predictor (i.e., independent variable) of the variation at hand: political identity, source-directed attitude and bias, language contact receptiveness, globalist/nationalist align-

ment, and mobility. This will provide a sense of how strongly loanword variation patterns with each factor. However, given the previous discussion about how these factors are not completely orthogonal to each other (§2.2), these factors should also be analyzed concurrently with each other to account for the possibility that one factor may exhibit this variation but only by inheritance from its correlation with another factor. For example, we may observe the variation of *Iran* and *Iraq* to pattern strongly both with political identity and with source-directed attitude when they are analyzed as individual predictors. However, when considered concurrently as predictors, we may observe that only source-directed attitude is needed to explain the variation at hand: i.e., once the effect of source-directed attitude is accounted for, political identity no longer explains any of the variation beyond that which was already explained by considering source-directed attitude.

Statistical analysis involved the generation of logistic mixed-effects models with predictors of interest and a step-up comparison of those models to test for significance (Winter 2014), with the procedure as follows. First, using the `lme4` package (Bates et al. 2014) in R (RCoreTeam 2015), a base logistic mixed-effects model of the data was generated using the `glmer()` function, including no independent variables of interest. This base model included a random intercept of word to account for the possibility that each word, across participants, might have a stronger default pronunciation. Then, for each independent variable of interest, a new model was generated that added only that variable as a fixed effect to the base model. Then, using R's `anova()` function, a Chi-square ANOVA test was used to determine whether the model incorporating the independent variable of interest showed a significantly improved fit of the data. Effect size and  $p$ -value of these ANOVA tests were used to rank independent variables by how strongly they contributed to model fit. The components of these individual-predictor models are those mentioned along with each variable when discussing the initial lone-predictor results. Then, step-up modeling was used to consider predictors concurrently. Any independent variable found to significantly improve model fit (if multiple, that with the strongest effect) was then used as a single fixed effect in a new base logistic mixed-effects

model, and the above cycle was repeated to test whether any remaining independent variable significantly improved model fit when additionally considered.

In the following statistical reporting, when discussing individual predictors independently, effect coefficients ( $\beta$ ), standard deviations ( $\sigma$ ),  $z$ -values ( $z$ ), and  $p$ -values ( $p$ ) come from the model including only the base model components and the single main effect of the social factor considered. When discussing the results of step-up modeling including multiple social factors concurrently, effect coefficients ( $\beta$ ), standard deviations ( $\sigma$ ), and  $z$ -values ( $z$ ) come from the final model including all fixed effects identified as significant contributors to model fit;  $p$ -values ( $p$ ) come from the Chi-square ANOVA test between that full model and a model excluding the effect of interest (Winter 2014). A more detailed documentation of the model formulae and the results of each step during the step-up comparison before reaching the final model are provided in the Appendices, with each respective appendix noted within the prose to follow.

### 2.3.2.3 Results

We'll start by examining just the variation of Arab-indexed *Iran* and *Iraq* to allow for the most direct comparison with previous studies of this variation (Hall-Lew et al. 2010; Silva et al. 2011). Figure 4 plots the data along the factor of political identity. An index score less than 0 means a more liberal/Democrat identity; a score greater than 0 means a more conservative/Republican identity. Looking across all participants (lefthand plot), political identity does seem to exhibit variation in the expected direction: Liberals are more likely to use the more source-like [a] variants than conservatives, and this is found significant ( $\beta=-0.657$ ,  $\sigma=0.204$ ,  $z=-3.229$ ,  $p=.0012$ ) in a lone-predictor logistic mixed-effects model. However, as mentioned above and apparent in the figure, a fuller spectrum of identity strength is better represented in the moderate-to-liberal half of the entire continuum. And, when only the data from those participants is examined (righthand plot), this effect is no longer significant. (See Table 3 for a summary of lone-predictor statistical models.) Therefore, while this



variation is apparent along the factor of political identity across broad political groups, it does not pattern as we might predict an act of political identity to: Those who identify more strongly as politically liberal do not show a significantly greater likelihood of using the more source-like pronunciations.

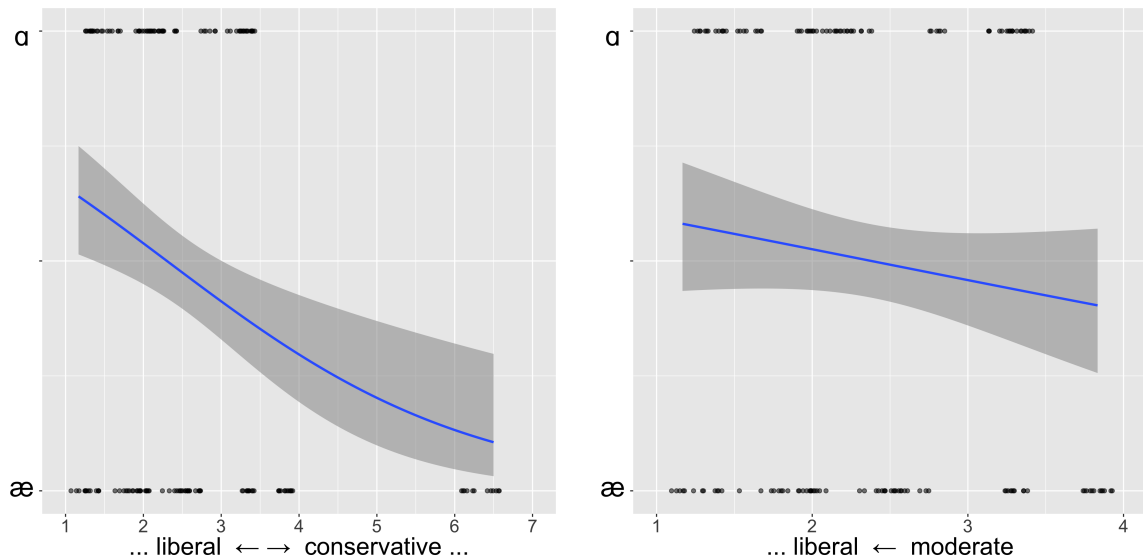


Figure 4: Political identity as predictor of  $\{Iran, Iraq\}$  variation

Variation of *Iran* and *Iraq* between  $[ɑ]$  and  $[æ]$  variants: the y-axis is the binary variable of which variant was used, the x-axis is the composite political identity score of the participant based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. The lefthand plot includes data from all participants; the righthand plot includes data only from those with a score of 4 or less (moderate- to liberal-identifying).

However, while degree of political identity strength does not capture within-group variation, such variation amongst those identifying as moderate to strongly liberal is still present. This is captured by participants' attitudes toward the source, as demonstrated in Figure 5. Participants who more strongly agree with statements expressing anti-Middle East sentiments in the Likert questionnaire (lefthand plot) are significantly more likely to use the less source-like  $[æ]$  variants. It was also predicted that this would be the case for participants with stronger negative implicit biases regarding Iran (as a narrowed substitute for the category of Arabs and the Middle East), if not more pronounced since participants might be less comfortable explicitly espousing such sentiments in the questionnaire. However, when examining the variation along this factor (righthand plot), the effect is not significant.

This difference suggests that loanword variation may not be as much a result of biased association regarding a source; rather, it may more accurately be considered a reflection of the attitudes that one explicitly identifies with and espouses.

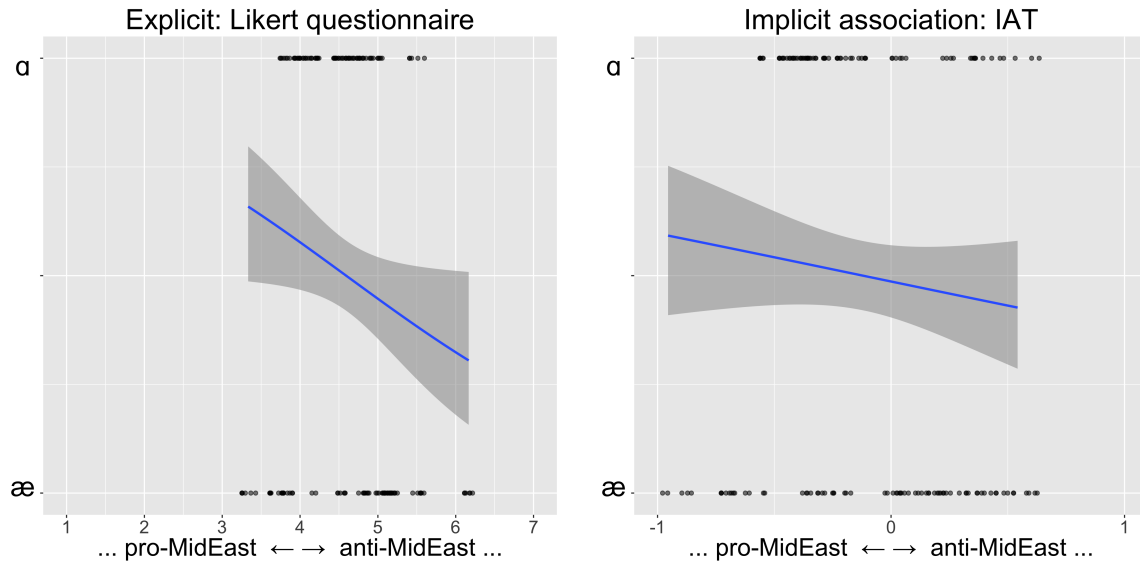


Figure 5: Middle East-directed attitudes as predictors of  $\{Iran, Iraq\}$  variation. Variation of *Iran* and *Iraq* between  $[\alpha]$  and  $[\ae]$  variants. The y-axis is the binary variable of which variant was used. In the lefthand plot, the x-axis is the composite place-directed attitude score of the participant based on their responses to the Likert questionnaire. In the righthand plot, the x-axis is the bias measurement as calculated with the algorithm provided by the Open IAT (Stafford and Scaife 2015), itself based off of the recommendations put forth by Greenwald et al. (2003): Here, a higher score means a stronger implicit bias associating Iran (as compared to Samoa) with negative attributes. Again, the line is a curved generalized linear model with a 95% confidence interval.

But, actually, the broader factors of language contact ideology and globalist/nationalist ideology seem to even more strongly capture this within-group variation. Figure 6 shows the results along each factor. One who is more globalist-aligning is significantly more likely to use the more source-like variants. Language contact ideology also patterns in the expected direction: One who aligns with a more receptive language contact ideology is also significantly more likely to use the more source-like variants.

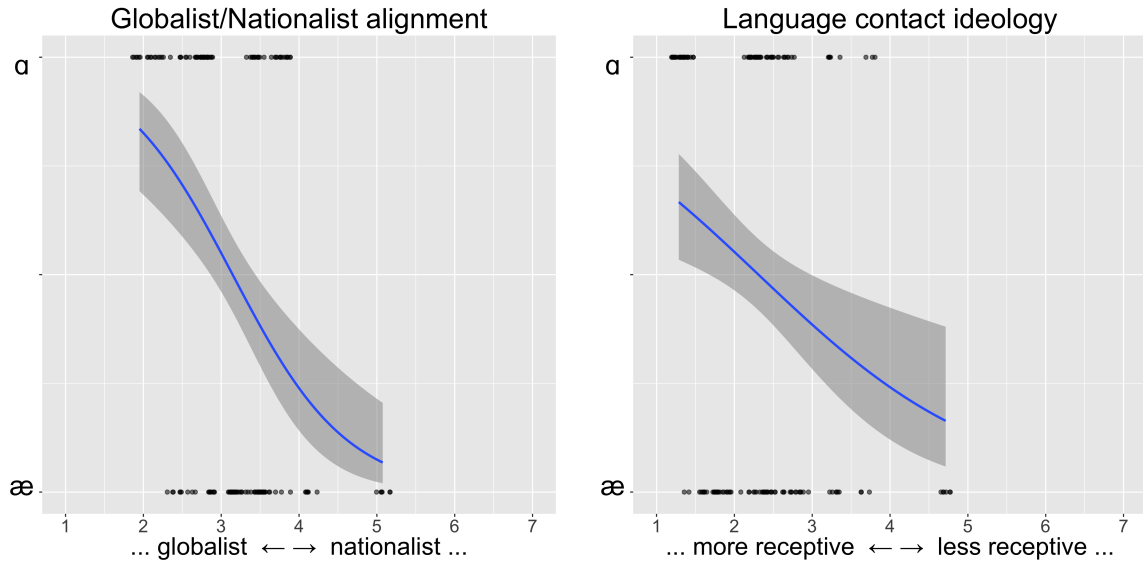


Figure 6: Ideologies as predictors of  $\{Iran, Iraq\}$  variation

Variation of *Iran* and *Iraq* between [ɑ] and [æ] variants: the y-axis is the binary variable of which variant was used, the x-axis is the composite globalist/nationalist alignment (lefthand) or language contact receptiveness (righthand) score of the participant based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval.

The pattern is similar when examining the variation across all loanwords targeted for analysis (including those from Table 1 that are not Arab-indexed). Political identity is only significant as a lone predictor when looking across groups ( $\beta=-0.358$ ,  $\sigma=0.094$ ,  $z=-3.816$ ,  $p=.00014$ ), but not within (see Table 3 below). And how strongly one aligns with a globalist ideology is the most consistent predictor. Additionally, while not thoroughly discussed or plotted above, it should be noted that participants' mobility (included as a factor in the questionnaire) was also examined and was not found to be a significant predictor of this variation. Table 3 provides a summary of the lone-predictor results for each factor, both the results specific to the variation of *Iran* and *Iraq* as well as those regarding the variation of the broader group of loanwords.

<i>{Iran, Iraq}</i> variation				
FACTOR	$\beta$	$\sigma$	$z$	$p$
globalist/nationalist alignment	-.921	.216	-4.26	$2.04e-5$
language contact ideology	-.551	.187	-2.956	.0031
Middle East-directed attitude (Likert)	-.346	.17	-2.036	.042
political identity	-.215	.166	-1.296	.195
Middle East-directed attitude (IAT)	-.181	.165	-1.095	.273
mobility	.091	.098	0.936	.349

Full set of target loans				
FACTOR	$\beta$	$\sigma$	$z$	$p$
globalist/nationalist alignment	-.475	.099	-4.758	$1.96e-6$
language contact ideology	-.296	.095	-3.11	.0019
Middle East-directed attitude (Likert)	-.212	.094	-2.257	.024
political identity	-.129	.093	-1.399	.162
mobility	.057	.055	1.036	.3
Middle East-directed attitude (IAT)	.014	.093	0.093	.88

Table 3: Lone-predictor statistical results

Results of logistic mixed-effects models including each factor of interest as a single independent variable, within the dataset of participants with a composite political identity score of 4 or lower (i.e., moderate- to liberal-identifying).

However, these factors are not necessarily orthogonal to each other. As argued above (§2.2), all of these factors are expected to correlate with political identity, and the variation observed along political identity may just be inherited by this correlation with another more direct predictor. Furthermore, *any* of these other factor pairings may also correlate with each other. Therefore, any factor may be inheriting this variation due its correlation with a stronger predictor. Table 4 provides a Pearson’s  $r$  score for the degree of correlation between each pair of the social factors considered. As is apparent, these factors do correlate with each other to some degree. For example, both those who align with a more political leaning and with a more receptive language contact ideology are also those who tend to align with a more globalist ideology. Therefore, examining these factors concurrently as predictors of the variation at hand can elucidate how much the variation observable along one factor may be explained by the variation observed along another.

	A	B	C	D	E	F
A political identity	1					
B Middle East-directed attitude (Likert)	.045	1				
C Middle East-directed attitude (IAT)	.189	-.067	1			
D language contact ideology	.486	.045	.174	1		
E globalist/nationalist alignment	.675	.122	.08	.715	1	
F mobility	-.216	.213	-.005	-.134	-.305	1

Table 4: Collinearity of social factors

Step-up modeling (as described above: §2.3.2.2) was used to concurrently analyze these individual factors alongside each other as predictors. Table 5 summarizes the fixed effects of the final model converged upon for both the entire set of variable words of interest as well as when looking only at the variation of *Iran* and *Iraq*.

<i>{Iran, Iraq}</i> variation				
FACTOR	$\beta$	$\sigma$	$z$	$p$
(intercept)	-.004	.179	-0.021	
globalist/nationalist alignment	-1.236	.273	-4.506	6.62 <i>e</i> -6
political identity	.467	.228	2.042	.041

Full set of target loans				
FACTOR	$\beta$	$\sigma$	$z$	$p$
(intercept)	-.263	.481	-0.546	
globalist/nationalist alignment	-.636	.128	-4.997	5.83 <i>e</i> -7
political identity	.253	.121	2.092	.038

Table 5: Step-up comparison results

Results of final logistic mixed-effects model after step-up comparison of independent variables as predictors, within the dataset of participants with a composite political identity score of 4 or lower (i.e., moderate- to liberal-identifying).

The results of this modeling suggest that the factor of globalist/nationalist alignment is the most consistent predictor, both when looking at the entire set of variable words of interest as well as when looking only at the variation of *Iran* and *Iraq*. The logistic mixed-effects model using only this fixed effect showed a significantly improved fit of the data. Only one other predictor, when included on top of this one, showed a significant improvement of model fit: political identity. However, this was in the reverse direction from that expected: Those more strongly aligning with a *liberal* political identity are those who are more likely

to use *less source-like* pronunciations once their globalist/nationalist alignment has been accounted for, hence the positive rather than negative direction of the effect ( $\beta$ ). Appendix B provides a more detailed summary of the step-up modeling performed.

#### 2.3.2.4 Interim discussion

These results again confirm that loanwords vary along political identity, at least across broad political groups and when considering political identity on its own. However, they provide a much deeper and more nuanced picture of this variation. For one, this variation does not appear to pattern as we might predict an act of political identity to: Strength of political identity *within* a political group does not significantly increase one's likelihood of using a particular variant in line with the across-group pattern; instead, this variation patterns in the reverse direction when looking at the strength of political identity within one political group. Furthermore, as hypothesized, factors more directly related to language contact are better predictors of loanword variation. Like the results of Study 1 above, those who identify as more politically moderate may align with a liberal ideology along the globalism/nationalism continuum but not align with a liberal political identity for other reasons, while their globalist/nationalist alignment seems to be a more consistent predictor of their loanword pronunciations.

Another notable difference in predictors was that between the methods of measuring participants' source-directed attitudes. Recall that explicit stance-taking regarding anti-/pro-Arab statements was identified as a better predictor of the variation of *Iraq* and *Iran* than implicitly measured associative bias. This runs counter to Babel's (2010) findings, which suggest that implicit bias is a stronger predictor of linguistic imitation. Part of the motivation for examining implicit bias in this study was an expectation that participants might feel less comfortable or willing to express certain biases explicitly. It was thought that one could have biases toward associating Arabs with negative attributes even if those biases are unfavorable in public opinion. Therefore, implicit bias examination was expected

to access and measure the attitudes that one might have but which they might not explicitly express, and which might still be reflected in their language.

However, as it turns out, this may be exactly why explicitly measured bias appears to be a better predictor. Loanword variation appears to pattern with the attitudes and ideologies that one is willing to express and identify with, in declaration and explicit support of such biases. In parallel, Nosek et al. (2007) find that Arab-directed attitudes show an uncommon asymmetry when measured implicitly and explicitly. In most cases (e.g., African-American-directed attitudes or age-related biases) they pattern together or implicit bias measurements show stronger effects than explicit measurements. However, Arab-directed attitude is found to show stronger effects when measured explicitly rather than implicitly: One might explicitly espouse negative attitudes regarding Arabs while not necessarily showing a strong processing bias in associating them with negative attributes. Nosek et al. also note that Arab-directed biases show one of the largest effects of political identity, both when measured implicitly and explicitly. This suggests that Arab-directed attitudes might be especially political and ideological: Alignment with a group promulgating such attitudes might override implicit biases, more so than one's implicit biases might guide their explicit behavior. The variation of Arab-associated loanwords patterning more strongly with explicit bias than implicit bias seems to reflect this. However, in line with the results above, where globalist/nationalist alignment is the strongest predictor and Arab-directed attitude is not identified as a significant conditioning factor alongside it, this variation may fall out from a broader globalist/nationalist ideological identity.

Rather than concern about one's willingness to explicitly align with certain attitudes and biases, the issue of willingness surfaces in a different way in this study. It was rather difficult to recruit conservative leaning participants. Recruitment in a metropolitan area and through the channels of a higher education institution likely contributed to a population asymmetry favoring liberal-identifying participants. However, this may be more than a sample's reflection of its population. Even when recruiting in a more conservative-populated

area, and even when commissioning a conservative with social ties to many more conservatives to help recruit, these efforts garnered only few such participants. This could be due to the recruitment method’s reliance on self-selection, while incentivized with remuneration. It seems that conservatives may be less likely to self-select as participants in such a study. Social psychology research of personality traits has found liberal-identifying participants to pattern with higher levels of Agreeableness (Caprara et al. 2003, 2006, Chirumbolo and Leone 2010)—suggesting a possibly higher willingness to cooperate—and Openness (Jost et al. 2003, Chirumbolo and Leone 2010)—suggesting less aversion to interacting with a stranger and/or taking part in an unknown set of tasks for that stranger. (This is also reflected in the patterns of interest in this study regarding ideology, with conservatives being less receptive to language contact and less globalist.) Other studies attentive to political identity, such as those of social and political psychology, corroborate this by also demonstrating a left-leaning asymmetry amongst participant pools (e.g., Cohen 2003, Okimoto and Gromet 2016). Therefore, this recruitment asymmetry itself has an interesting relation to the questions at hand, reflecting some of the expected patterns in a different way. This motivates future research to consider how to improve the balance of political identities represented within the participant pool, while still avoiding the potential issues that might arise by overtly recruiting around political identity.

Limitations such as that of political identity representation motivate follow-up Study 3 (§2.3.3), which is discussed next. For one, as just mentioned, most participants identified as either moderate or liberal-leaning in their political identity. While a spectrum of the *strength* of that identity was captured and measured, we would ideally be able to see data representing a wider spectrum across political identities. Furthermore, a larger sample size is motivated in comparing the predictors under study, which are collinear to some degree. And, as also discussed regarding Study 1 (§2.3.1), members of this study’s participant pool identified as being from many different parts of the US. More tightly controlling for US region is also ideal, to ensure that the social factors of primary interest do not conflate with



it. The following study is an attempt at overcoming such limitations.

### **2.3.3 Study 3: Increasing political identity representation; controlling region**

This study is a follow-up version of Study 2 (§2.3.2), with some methodological modifications intended to overcome the limitations identified above. The hypotheses, however, stay the same: It is expected that political identity will exhibit the same variation of established loanwords but that the other social factors of interest will be stronger predictors of this variation by nature of their more direct connection to language contact.

#### **2.3.3.1 Methods**

One methodological modification was that of population sampling. Instead of New York City, the study was conducted in Bloomington, Indiana, which served multiple purposes. Political polling data and election results suggest that the population of Bloomington is much more balanced along political identity than the population of New York City (*Politico* 2016). A recruitment method of remunerated self-selection via flyer advertisements and online postings (as in Study 2) was still used, but this did successfully result in a wider distribution of political identities amongst participants, extending further into the conservative-identifying half of the spectrum with a fairly balanced distribution (discussed further below). This also successfully resulted in participants' US region being more tightly controlled, with a large majority of participants identifying their hometown as a city in Indiana or a closely neighboring city in Illinois.

Some changes were also made regarding the methods of elicitation and analysis. Elicitation included more variable loanwords, especially with the intention of including non-placename loanwords. With a concern that participants in Study 2 may have been too attentive to their pronunciations, elicitation was no longer a wordlist task but a sentence reading task. Words were embedded in sentences designed to feel as politically uncharged and semantically generic as possible: e.g., “The geography of Iraq varies considerably from

north to south.” As in Study 2, filler stimuli were also included amongst the target stimuli of interest, and this was also following a task of listening to short stories and reading short sequels aloud (to be discussed in Chapter 4).

The analysis of variable pronunciations was also done with more fine-grained attention to words that may contain multiple loci of variation. For example, *Chile* is variable along two dimensions: the first syllable nucleus ([i]~[ɪ]) and the second ([e]~[i]). As another example, *foyer* is variable along three dimensions: stress placement (final~initial), rhotic presence (non-rhotic~rhotic), and final vowel presence ([e,ɛɪ]~[ɪ]). This is somewhat inspired by Poplack et al.’s (1988) calculation of an ‘index of phonological integration’ when examining the variation of English-to-French borrowings, but it differs in a crucial way. Poplack et al. coded every segment of every word as either a French-like production or English-like production (though noting cases of ambiguity/overlap between the two languages’ sound systems). Here, only segments or features that did indeed exhibit variation in the data at hand were considered part of the envelope of variation. (How outliers and rare variants were identified and dealt with is briefly discussed in the next section: §2.3.3.2.) Table 6 lists the variable loanwords of interest, along with the identified locus/loci of variation within each and how variants were coded.

<u>brus</u> chetta	[sk]	~	[ʃ]	<u>Nev</u> ada	[ɑ]	~	[æ]
<u>Ch</u> ile	[i],[e]	~	[ɪ],[i]	<u>Par</u> aguay	[ɑ],[aɪ]	~	[æ],[e]
<u>fo</u> yer	[σ'σ],[e,ɛ],[Ø]	~	['σσ],[Ø],[ɪ]	<u>Que</u> bec	[k]	~	[kw]
<u>g</u> enre	[ʒ]	~	[ʒ]	<u>s</u> piel	[ʃ]	~	[s]
<u>Haw</u> aii	[ʔi]	~	[i]	<u>tama</u> les	[e],[s]	~	[i],[z]
<u>I</u> ran	[ɪ],[ɑ]	~	[aɪ],[æ]	<u>Tok</u> yo	[j]	~	[i]
<u>I</u> raq	[ɪ],[ɑ]	~	[aɪ],[æ]	<u>tsu</u> unami	[ts]	~	[s]
<u>Mu</u> slim	[u,ʊ],[s],[ɪ]	~	[ʌ],[z],[ə]				

Table 6: Variable words of interest: more source-like ~ less source-like  
 These are the words of interest that were elicited in the sentence reading task. The parts underlined are those used in the coding of utterances between more source-like (left) and less source-like (right) pronunciations, with each locus of variation and its respective codings represented.

Minor changes were also made to the following IAT and Likert questionnaire tasks. For the IAT, instead of comparing implicit biases regarding Iran and Samoa, Iraq and Indonesia

were substituted in their place. This was to match the different places already previously mentioned to participants in the preceding imitation task (to be discussed in Chapter 4), which was modified between pilot and post-pilot versions. The place-directed attitude factor in the Likert questionnaire was similarly modified, referring to Iraq and Indonesia specifically rather than broader regions (like in the pilot: ‘Middle East’ and ‘Polynesia’). The rest of the questionnaire mostly remained the same. However, a new set of statements was introduced to examine how strongly one aligns with a ‘prescriptivist’ or ‘descriptivist’ ideology about language, both regarding native-speaker variation and non-native accentedness. These are provided in Table 7.

FACTOR	STATEMENT
Prescriptivism	
descriptivist	I enjoy hearing different regional accents of English. People with strong foreign accents are just as likely to be smart as people without accents.
prescriptivist	It’s annoying when someone doesn’t use proper English grammar. Customer service representatives shouldn’t have strong foreign accents.

Table 7: Likert agreement statements: prescriptivism

### 2.3.3.2 Results

Recruitment garnered 98 participants. Compared to Study 2, a much wider distribution of political identity was represented in this participant pool: a bimodal distribution with each mode on either side of the halfway point of the continuum, as apparent in Figure 7. It is notable that this entire distribution is shifted left of center. However, this may be inevitable, given that the generation sampled (Millenials) is shown to significantly favor identifying as Democrat and holding consistently liberal views (*Pew Research Center* staff 2018). It may also be an artifact of the time at which this study was conducted: early spring of 2017. *Pew Research Center* (2017) reports that during the first four months of 2017, immediately following the 2016 election of Republican President Donald Trump, there was a drop in the favorability ratings of the Republican Party (both when measured cardinally

and in relation to the favorability ratings of the Democratic Party). The same study finds that, even amongst the Republican Party’s own members, there is a more widely shared feeling that the Republican Party is divided. This could explain why, even though there were participants who identified as conservatives, fewer felt comfortable identifying as such in strong, absolute terms.

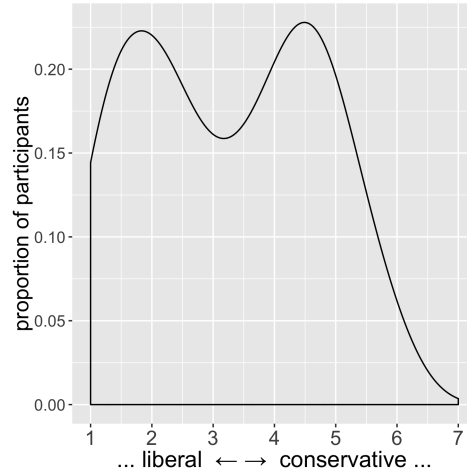


Figure 7: Political identity distribution of participants

The distribution of all participants’ composite index scores for the factor of political identity, based on responses to the 7-point Likert questionnaire.

Political identity also appears to have been teased apart, at least to some degree, from the other social factors of interest. Table 8 provides a Pearson’s  $r$  score for the degree of correlation between political identity and each other variable. Given that the maximum value is 1 (either positive or negative), the largest values being below  $|0.4|$  suggest that a good amount of the data does not conflate political identity with these other factors. However, the two strongest correlations are in line with the previous qualitative analysis and discussion (§2.2). Political identity patterns in the expected direction with language contact ideology and alignment with a globalist/nationalist ideology: Conservatives tend to be less receptive to language contact and align with a more nationalist ideology. With this in mind, we will proceed to examine the data in a similar fashion as in the previous study, first examining the social factors’ strength as lone predictors and then using step-up model comparisons to

identify which best explain the variation at hand when considered concurrently.<sup>3</sup>

	FACTOR	CONTINUUM		<i>r</i>	
	Language contact ideology:	more receptive	~	less receptive	0.382
	Global/Nationalist ideology:	globalist	~	nationalist	0.329
	Prescriptivism:	descriptivist	~	prescriptivist	0.223
	Mobility:	more mobile	~	less mobile	-0.148
	Place-directed attitude:	Iraq favoring	~	Iraq disfavoring	0.068
	Implicit association:	Iraq favoring	~	Iraq disfavoring	-0.003

Table 8: Collinearity with political identity

Pearson’s *r* scores of correlation between political identity and each other social factor of interest. Scores are ordered by correlation strength, with values closer to 0 meaning weaker correlations while maximum scores are 1 and -1. Positive scores mean that the righthand end of the continuum more strongly correlates with conservative political identity, and the reverse for negative scores.

First, as in Study 2, the variation of only *Iran* and *Iraq* is examined to allow for more direct comparison with previous studies (Hall-Lew et al. 2010; Silva et al. 2011). Figure 8 plots the utterances of *Iran* and *Iraq* across the fuller spectrum of political identity represented in this study. Dot sizes in the plot represent the number of data points which would otherwise be overlapping due to multiple participants having the same composite index score for political identity. The y-axis represents the ‘number of divergences’ from the source form, in which each use of the less source-like pronunciation of a variable within the word was counted as 1 point: e.g., [ɪˈɪɾk] = 0, [ɪˈɪæk] = 1, [aɪˈɪæk] = 2. (The case of one using [aɪ] for the first syllable and [a] for the second was observed infrequently, but a handful of times across both words and multiple participants.) As is clear from the data, there is a strong preference across the board for the [aɪ...æ...] pronunciations, confirming that this is a well-established adaptation. (There are notably two utterances, from different speakers, that received a score of 3 because of a rare shifting of stress to the first syllable: e.g., [ˈaɪɪæk].) It is also apparent that the data patterns with political identity in the direction that would be expected: More conservative-identifying participants significantly prefer the less source-like

<sup>3</sup>Statistical analysis was performed in a fashion similar to that described above (§2.3.2.2). However, given that the dependent variable is now non-binary, the mixed-effects models were linear rather than logistic. When discussing the results of step-up predictor comparison, effect coefficients ( $\beta$ ), standard deviations ( $\sigma$ ), and *t*-values (*t*) come from the final model including all fixed effects identified as significant contributors to model fit; *p*-values (*p*) come from a Chi-square ANOVA test between that full model and a model excluding the effect of interest (Winter 2014).

pronunciations. However, once again, the stronger predictor is participants' agreement with statements aligning with more globalist or more nationalist ideologies, as shown in Figure 9.

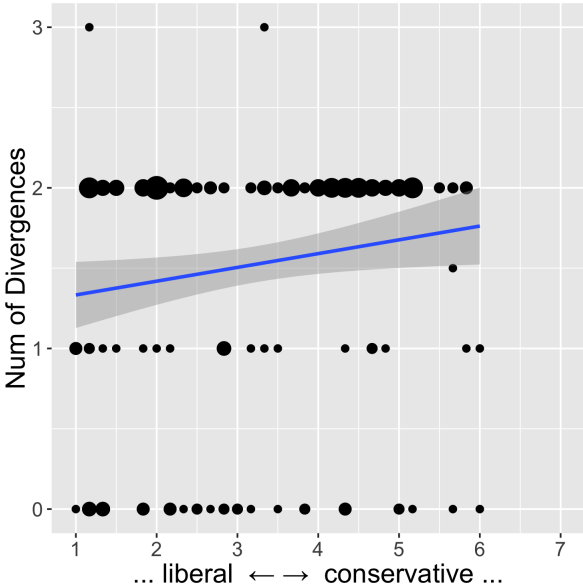


Figure 8: Political identity as predictor of  $\{Iran, Iraq\}$  variation

Variation of *Iran* and *Iraq*: the y-axis represents the number of variables in which an utterance used a less source-like variant, the x-axis is the composite political identity score of the participant based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Dot sizes represent the number of data points that would otherwise be overlapping.

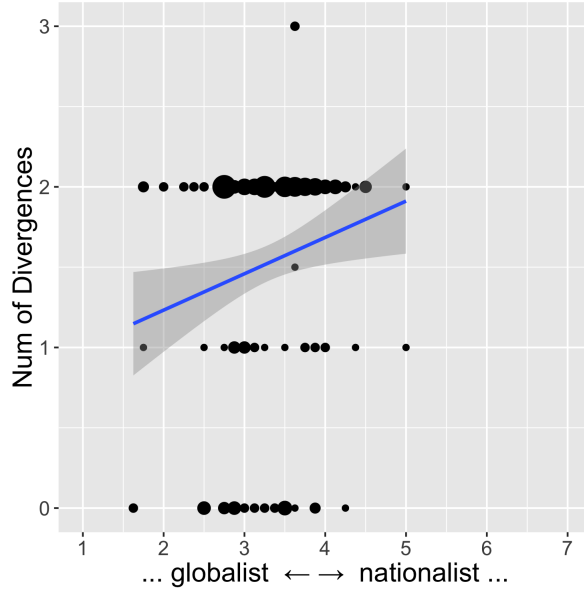


Figure 9: Globalist/Nationalist alignment as predictor of  $\{Iran, Iraq\}$  variation  
 Variation of *Iran* and *Iraq*: the y-axis represents the number of variables in which an utterance used a less source-like variant, the x-axis is the composite globalist/nationalist score of the participant based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Dot sizes represent the number of data points that would otherwise be overlapping.

As with the previous dataset analyzed, the social factors of interest are not assumed to be orthogonal and perfectly teased apart from each other. Table 9 provides a Pearson’s  $r$  score for the degree of correlation between each pair of the social factors considered.

	A	B	C	D	E	F	G
A political identity	1						
B Middle East-directed attitude (Likert)	.068	1					
C Middle East-directed attitude (IAT)	-.003	.094	1				
D language contact ideology	.382	.307	-.082	1			
E globalist/nationalist alignment	.329	.114	-.059	.51	1		
F mobility	-.148	-.24	.062	.011	.117	1	
G prescriptivism	.223	.196	.011	.459	.323	-.054	1

Table 9: Collinearity of social factors

Given this collinearity, we turn again to a step-up statistical analysis to compare predictors alongside each other: base mixed-effects model (in this case linear rather than logistic, given the dependent variable); model with a single factor of interest included as a fixed effect; Chi-square ANOVA test for significant improvement of model fit; repeat

after keeping the strongest (if any) fixed effect found to significantly improve model fit. The number of divergences was treated as the dependent variable and, like in Study 2, a random intercept was specified per word. Table 10 provides the results of the first cycle: lone-predictor models for each independent variable. Table 11 provides the results of the final model converged upon after step-up comparison of independent variables. Appendix C provides a more detailed summary of the step-up modeling performed.

FACTOR	$\beta$	$\sigma$	$t$	$p$
globalist/nationalist alignment	.139	.057	2.472	.014
political identity	.125	.057	2.205	.028
mobility	.097	.057	1.701	.089
prescriptivism	.073	.057	1.287	.197
language contact ideology	.065	.057	1.142	.252
Iraq-directed attitude (IAT)	-.034	.057	-0.587	.555
Iraq-directed attitude (Likert)	.007	.057	0.121	.903

Table 10: Lone-predictor statistical results:  $\{Iran, Iraq\}$  variation  
Results of linear mixed-effects models including each factor of interest as a single independent variable.

FACTOR	$\beta$	$\sigma$	$t$	$p$
(intercept)	1.525	.056	27.218	
globalist/nationalist alignment	.181	.059	3.062	.0023
prescriptivism	.129	.059	2.207	.027

Table 11: Step-up comparison results:  $\{Iran, Iraq\}$  variation  
Results of final linear mixed-effects model after step-up comparison of independent variables as predictors.

A first cycle identifies globalist/nationalist alignment as the strongest lone predictor. While political identity is also identified as a significant predictor when examined as the only fixed effect (see Table 10), it is no longer identified as such in a subsequent model that considers both political identity and the factor of globalist/nationalist alignment concurrently ( $\beta=0.089$ ,  $\sigma=0.059$ ,  $t=1.508$ ,  $p=.129$ ). This suggests that, while political identity exhibits this variation to a strong degree, the observable political variation (and more) is accounted for by considering globalist/nationalist alignment instead.

There is, however, one final factor of interest identified as a significant predictor alongside globalist/nationalist alignment in the same model: prescriptivism. As shown in Figure



10, those who identified as more prescriptivist showed a stronger preference for the more source-like variants. This is in line with Boberg’s (1999) conclusion that the [ɑ] variant of ‘foreign (a)’ carries greater linguistic security: Speakers who hold stronger feelings that they should use “correct” language are more likely to use a more source-like variant. The final model including only factors significantly improving fit of the data includes only these two factors: globalist/nationalist alignment and prescriptivism. However, political identity is still notably identified as a trend when considered as an additional factor to incorporate ( $\beta=0.107$ ,  $\sigma=0.059$ ,  $t=1.802$ ,  $p=.069$ ).

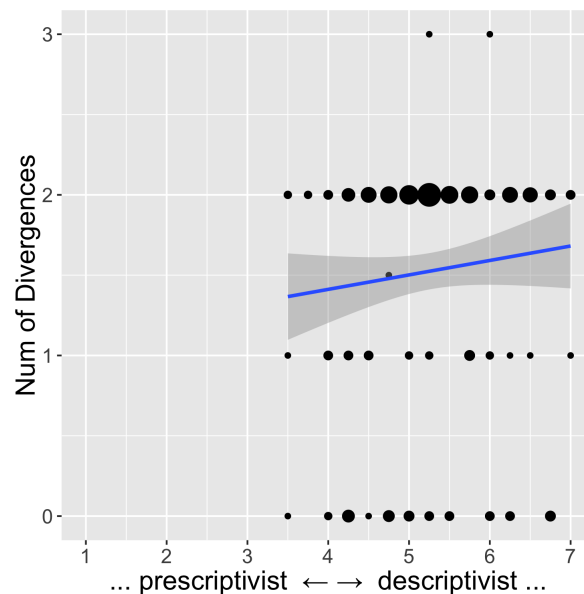


Figure 10: Prescriptivism as predictor of  $\{Iran, Iraq\}$  variation

Variation of *Iran* and *Iraq*: the y-axis represents the number of variables in which an utterance used a less source-like variant, the x-axis is the composite prescriptivism score of the participant based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Dot sizes represent the number of data points that would otherwise be overlapping.

Expanding the analysis to the entire set of loanwords of interest (provided above in Table 6), a similar pattern is apparent. Political identity does exhibit variation in the expected direction, with more liberal-identifying participants more frequently using more source-like pronunciations of loanwords; but, the effect of how globalist/nationalist a participant identifies appears to be stronger. This is shown in Figure 11, which is similar to the previous figures but with the y-axis converted proportionally to a scale of 0-1 to account for the fact that

some words had more loci of variation than others, like the ‘index of phonological integration’ used by Poplack et al. (1988). Some outlier pronunciations were identified, suggesting that the speaker may not have recognized the word or been familiar with it beforehand (e.g., [ɟə'neɪ] for *genre*, [ˈpæɪəge] for *Paraguay*), and therefore excluded.

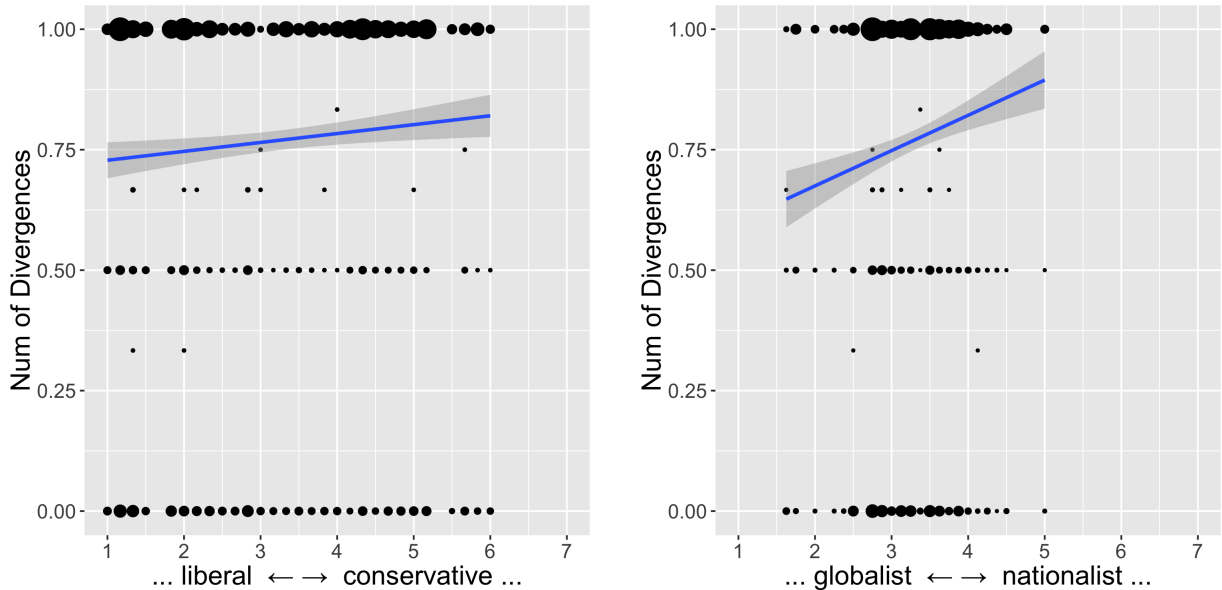


Figure 11: Political identity, globalist/nationalist alignment as predictors of general loanword variation

Variation of the full loanword set of interest: the y-axis represents the number of variables in which an utterance used a less source-like variant divided by the number of loci of variation identified for that word, the x-axis is the composite political identity score (lefthand) or globalist/nationalist score (righthand) of the participant based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Dot sizes represent the number of data points that would otherwise be overlapping.

Step-up statistical modeling was again performed. Unlike in the above figure, the dependent variable was still a tally of the number of divergences from the source form. Normalization across words was achieved, like in the previous statistical modeling, by including a random intercept per word in every model. Unlike the set of *Iran* and *Iraq*, though, it was also considered that this full set of words covers a much more diverse array of sources, degrees of establishment, and sound variables. Therefore, each cycle of modeling included the consideration of a random *slope* per word as well. That is, a word like *genre* may show a stronger default toward its more source-like ([ʒ]) variant than *Iran* or *Iraq* might (motivating

a different intercept); but, there might also be a different effect strength per word, such as the variation of *genre* not being as strongly predicted by one’s political identity as the variation of *Iraq* might be (motivating a different slope). Table 12 provides the results of the first cycle: lone-predictor models for each independent variable. Table 13 provides the results of the final model converged upon after step-up comparison of independent variables. And, again, C provides a more detailed summary of the step-up modeling performed.

FACTOR	$\beta$	$\sigma$	$t$	$p$
globalist/nationalist alignment	.069	.014	4.907	1.02 e-6
mobility	.037	.014	2.621	.009
language contact ideology	.036	.014	2.565	.01
political identity	.054	.019	2.771	.01
prescriptivism	-.005	.014	-0.326	.744
Iraq-directed attitude (Likert)	-.004	.014	-0.259	.796
Iraq-directed attitude (IAT)	-.0006	.014	-0.046	.964

Table 12: Lone-predictor statistical results: All target loans  
Results of linear mixed-effects models including each factor of interest as a single independent variable.

FACTOR	$\beta$	$\sigma$	$t$	$p$
(intercept)	1.204	.163	7.39	
globalist/nationalist alignment	.049	.015	3.356	.0008
mobility	.038	.014	2.65	.008
political identity	.044	.02	2.168	.036

Table 13: Step-up comparison results: All target loans  
Results of final linear mixed-effects model after step-up comparison of independent variables as predictors.

Results again identify globalist/nationalist alignment as the strongest predictor, in the expected and previously observed direction. The next factor identified as a significant predictor is participants’ mobility, with those identifying as more mobile being those more likely to use more source-like pronunciations. Finally, in this case, political identity itself is also identified as a significant predictor in the expected direction: Unlike in Study 2, those who are more liberal-identifying are those who are more likely to use more source-like pronunciations. And, while the factor of prescriptivism is not identified as a significant predictor like it was regarding just the variation of *Iran* and *Iraq*, it is identified as a trend

( $\beta=0.025$ ,  $\sigma=0.015$ ,  $t=1.697$ ,  $p=.089$ ) in the same direction. Figure 12 shows the data plotted along the additional factors of mobility and prescriptivism. As is especially notable in the plot regarding prescriptivism, the effect does not appear very strong; however, it is important to remember that these plots consider only their respective factors as independent variables, unlike the multivariate statistical models.

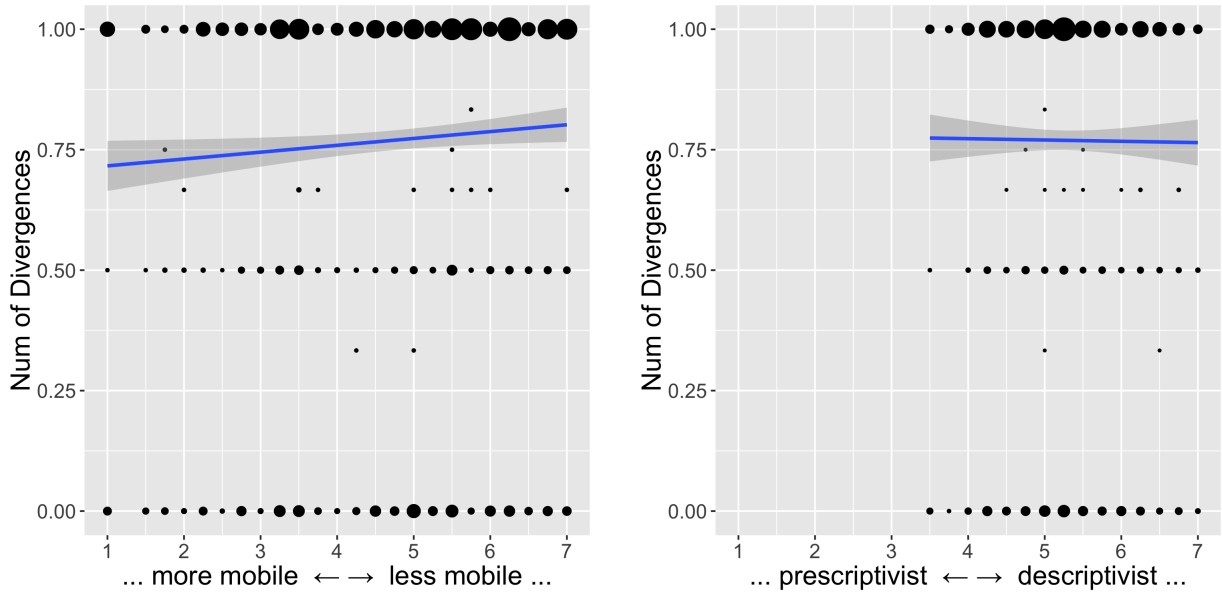


Figure 12: Mobility, prescriptivism as predictors of general loanword variation

Variation of the full loanword set of interest: the y-axis represents the number of variables in which an utterance used a less source-like variant divided by the number of loci of variation identified for that word, the x-axis is the composite mobility score (lefthand) or prescriptivism score (righthand) of the participant based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Dot sizes represent the number of data points that would otherwise be overlapping.

One random slope per word was also motivated in the above modeling: that of the effect of political identity. This suggests that, while the effects of globalism/nationalism, mobility, and prescriptivism appear to hold fairly consistently across the diverse set of loanwords, the effect of political identity is stronger for certain words compared to others. Table 14 lists each word, along with its intercept and assigned slope for the effect of political identity. First it is apparent that, for all words but one, the intercept is above half its maximum number of divergences: All words show a default usage of their maximally adapted variants, with the sole exception of *genre* exhibiting a stronger default preference for the more source-

like [ʒ] variant. Regarding slopes (by which the results in Table 14 are ordered), all words that are Arab-indexed or associated with Latin America are above all others, showing the strongest political divide in the same direction. (The only Spanish word that is an exception to this, *Nevada*, is that which pertains to a US state.) The fact that these words show the strongest political effect is not counterintuitive when considering how, as demonstrated above, attitudes toward Arabs, Arab-Americans, and Muslims show an especially strong political divide (Arab American Institute 2014; Nosek et al. 2007), as well as attitudes toward those who identify as Latinx and the use of the Spanish language in the US (Barker et al. 2001; Santoro 1999; de la Garza and DeSipio 1994; Mindiola and Gutierrez 1988).

WORD	INTERCEPT	SLOPE	WORD	INTERCEPT	SLOPE
Muslim	2.76	0.141	Hawaii	0.97	0.024
Iraq	1.56	0.098	bruschetta	0.93	0.02
Chile	1.2	0.094	spiel	0.67	0.018
tamales	1.69	0.074	tsunami	0.97	0.011
Paraguay	1.72	0.07	Nevada	0.72	0.009
Iran	1.46	0.069	Quebec	0.81	0.006
foyer	1.77	0.055	genre	0.34	-0.005
			Tokyo	0.52	-0.018

Table 14: Word-specific intercepts and political identity slopes

Each variable word of interest along with its intercept (in a sense, the modeled average number of divergences from the source form) and slope assigned for the effect of political identity. Words are ordered by slope, with the lefthand group consisting of those with stronger effects of political identity.

## 2.4 Discussion and Conclusions

The results of this study, in many ways, confirm the hypotheses theoretically motivated and set out to be tested. The variation of loanwords between more and less source-like pronunciations across political identities, as found by Hall-Lew et al. (2010) in US politicians’ pronunciation of *Ir*[ɑ~æ]*q*, extends to the speech of non-politicians. This pattern also extends to the broader class of loanwords, including those of other sources and other sound variables: e.g., *Chil*[e~i], [k~kw]*ebec*. Furthermore, when considering political identity alongside other social factors that appear to also be relevant to loanword variation, and which also correlate with political identity, a picture emerges that suggests the political vari-

ation of loanwords may be second-order. One's alignment with a more globalist ideology is a stronger predictor that they will use more source-like loanword pronunciations, whether or not they also align with a liberal political identity.

These conclusions require some nuanced attention, which will be further raised and discussed in this section. However, there are multiple broad takeaways and implications. For one, the observations made here shine further light on our understanding of US politics. The variation of loanwords along political identity appears to fall out from the attitudes and ideologies that pattern *with* US politics, which therefore corroborates this relationship and demonstrates how language reflects it. Yet, while the *relationship* of attitudes and ideologies with political identity is corroborated in some ways, the observations made here simultaneously reveal that these things are not one and the same: Political identity may correlate with globalist/nationalist alignment; but, there appear to be those who align with the other end of this particular ideological spectrum more strongly than their peers who might, for other reasons, still share the same political identity.

These observations also have implications for future sociolinguistic research, suggesting that loanwords as a broad class (from multiple sources) can co-vary with significant social meaning—a social meaning that, itself, regards one's ideology about the broader world and their relationship to it. Much previous work regarding loanword adaptation (as reviewed in §2.2) has suggested that identity, attitude, and intensity of language contact can influence its outcomes, especially resulting in the sociolinguistic variation of loanwords. But the current study further and in a more directly empirical method supports the interpretations put forth by Thomason (2001) and Poplack et al. (1988) that speakers' ideologies can influence such outcomes as well. And, while the ideology put forth has often been that regarding language contact in particular, the observations made here suggest that this may only be a part or correlate of an even broader ideology of globalism. Even the effect of source-directed attitude is suggested to fall out from this, with globalist/nationalist alignment faring better as a predictor of the variation of Arab-indexed loans than Arab-directed attitudes and biases.

This motivates further work examining loanword variation in American English, a prime case in which to examine such broad effects that may apply across loanwords of various sources. Much language contact-related research involving English considers contexts in which it is a dominant, superstrate language and examines how other languages and/or communities navigate and exhibit the effects of this contact. But, English still shows loanword manifestations of this low-intensity (on the English end) contact with multiple languages. This allows us to examine factors which may be more broad and less particular to the contact language or context, like general language contact ideology and globalism. And this study's results suggest that such factors indeed play a role in the outcomes of language contact and need to be considered and examined further.

A more nuanced discussion of how political identity conditions loanword variation is warranted. The results of Study 2 (§2.3.2) and the variation of *Iran* and *Iraq* in Study 3 (§2.3.3) suggest that political identity is considered a significant predictor only when examined on its own but not when considered alongside globalist/nationalist alignment. However, political identity does emerge as a significant predictor when looking across the broad suite of loanwords in Study 3, though still a weaker predictor than globalist/nationalist alignment. This suggests that, while the political indexicality of loanword variation may be a result of its correlation with other predictors, it may not be a mere byproduct of that correlation, per se. Silverstein's (2003) introduction of 'indexical order' and Eckert's discussion of 'indexical fields' (2008) allow us to consider how a language variable may have multiple socially related indexicalities without forcing us to identify only one such indexicality as true and the rest as byproducts. The indexicality of a language variable can be fluid, multiplicitous, and subject to reconstrual due to the connectedness (whether by sociological correlation or ideological association) of social identities and attributes.

As seems to be the case here, the variation of loanwords may be more primarily predicted by globalist/nationalist alignment; but, the fact that this alignment is correlated and associated with political identity opens the door for loanword variation to be reconstrued

as a signifier of political identity. And, the fact that political identity is especially salient and potent in the US zeitgeist (Abramowitz 2013; Westfall et al. 2015) likely predisposes this indexicality to be gravitated toward in such reconstrual. While the apparent reconstrual in this case does not appear to be absolute (i.e., a complete shift from the globalist/nationalist indexicality), it does appear that loanword variation has gained a second-order political indexicality on top of that which is more primary. Furthermore, given the identification of a random slope of the political identity effect per word (§2.3.3.2), certain words appear to have gained this indexicality more strongly than others. The variation of words associated with Arabs, Islam, and Latin America shows the strongest political indexicality, parallel to how current US discourse around related topics shows an especially high degree of political charge.

There are further indexicalities that these findings also highlight. As discussed when theoretically motivating the social factors to be considered in this analysis (§2.2.3), one's alignment with a globalist ideology could also relate to one's mobility and, further still, to one's socioeconomic status (due to the resources that mobility may require) resulting in an indexicality of prestige. In the results of Study 3, one's mobility is also identified as a significant predictor: One with more mobility is more likely to use more source-like pronunciations of loanwords. This could reflect that those who are more mobile may also be those who are more likely to be familiar with the more source-like (and less established) pronunciation variants of loanwords, such as by having had more opportunity to come into contact with other languages or language varieties. However, it still appears that one with a more globalist alignment, even if they might not be especially mobile, will also be more likely to use such pronunciations. We also see an effect of prescriptivism (significant in the variation of *Iran* and *Iraq*; a trend when looking across all loanwords examined): One with a stronger feeling that people should use "correct" language is more likely to use more source-like pronunciations of loanwords. This suggests that more source-like pronunciations carry greater linguistic security (in line with Boberg's [1999] findings). The combination



and intersection of these indexations—globalism, mobility, and linguistic security—suggests that the use of more source-like pronunciations may carry capital in what we might think of as the global linguistic market (Bourdieu 1977; Piller 2001; Zhang 2005; Blommaert 2010). This is further addressed in the following study (Chapter 3) examining how listeners index and evaluate loanword variation upon hearing it.

We have considered how this variation patterns, furthering our understanding of its indexicality and how additional, related indexicalities appear to have arisen. Besides its *indexicality*, though, we should also consider how the *variation* itself has arisen. Even when controlling for region in Study 3, substantial variation is observed. This could be attributable to input and acquisition. It is possible that one simply learns one variant from their family or close community and that, in parallel, they inherit similar ideologies and identities. However, this consideration merely pushes the question back a generation. How did these variants come to pattern with relevant social differences in the first place?

Two notions, which are not mutually exclusive, may be useful in addressing this: change during the lifespan and social influences on initial loanword adaptation. As noted above, even those who may not be particularly mobile but who nonetheless align with a more globalist ideology are more likely to use more source-like pronunciations. It is possible that these speakers have changed their pronunciations: that they grew up using the more established adaptations but, upon encountering pronunciations that they identified (or conjectured) as more closely resembling the source forms, they were inclined to prefer these pronunciations as a kind of deference to those who more strongly associate with those words or languages, in line with a more globalist ideology. Boberg's (1997) findings that the [ɑ] variant of 'foreign (a)' is becoming more prevalent amongst younger generations of speakers corroborate this notion. Boberg's findings suggest that one might change to the [ɑ] variant in spite of their preceding generations using the [æ] variant, and the current observations shed further light on who might be especially likely to do so. This does not necessarily imply an *agentive* decision above the level of consciousness to change one's pronunciation; but, it also

does not preclude the possibility. Those who are more globalist might also be more likely to agentively seek out the source forms of foreign words they use in a similar kind of deference, just as one might look up how a new correspondent's name is spelled before addressing a correspondence to them as a way to avoid the chance of losing favor with them by misspelling it (which would also likely correlate with how strongly one defers to that correspondent).

In addition to social factors influencing whether one might change their pronunciation of a loanword already borrowed into the language, the same factors may mediate how strongly one replicates or adapts the form of a new loanword entering the language. As suggested above, one who is more globalist might be more likely to agentively seek out the source pronunciation of a loanword with an intent to replicate it as closely as possible. But, such a factor could also mediate how strongly one replicates the pronunciation of a new foreign word they've encountered passively. This could therefore explain how the variation between more and less source-like pronunciation comes about in the first place: Those who are more globalist do not just acquire or adopt the more source-like variants of already established loanwords; such variants may be available because those who are more globalist were more attentive (not necessarily agentively) to replicating the source form as close as possible, starting from the very beginning of the path of adaptation and establishment.

This distinction between established vs. new loanwords could also account for residuals in the variation of established loanwords: i.e., cases where a speaker might have been predicted to use one variant based on their identities and ideologies but used another. Such discrepancies could be due to prior exposure. One may prefer to use a particular variant because that was their sole/majority input when first learning the word; they may be resistant to changing their pronunciation, and/or they may prefer matching the language usage of those around them. However, when the factor of prior exposure is stripped away, these social effects may still emerge. The study in Chapter 4 is intended to examine this potential for initial loanword adaptation to be socially influenced. As mentioned in the Methods descriptions of Studies 2 (§2.3.2.1) and 3 (§2.3.3.1), the task eliciting the utterances of established

words was preceded by a story-reading task. Participants hear nonce words in short stories and, after exposure, they read aloud short sequels that elicit utterances of the same nonce words. In some cases, participants are led to think the nonce words are loans; in other cases, unfamiliar English words. Then, the same social factors are analyzed for their influence on how strongly participants replicate the exposure form, following a phonetic imitation experimental paradigm (e.g., Babel 2009; 2010; Yu et al. 2013; Lev-Ari and Peperkamp 2014). This approach therefore examines how social factors influence initial adaptation, removing the factor of prior exposure.

## Chapter 3

### Perceptual indexation and evaluation of loanword variation

#### 3.1 Introduction: Considering perception

In Chapter 2, the variation of established loanwords between more and less source-like pronunciations was observed to pattern significantly along multiple social dimensions. In particular, and seemingly the strongest and most core predictor, those who are more globalist-aligning are those who use more source-like pronunciations. Loanword variation was also observed to pattern with other social factors, however, in line with previous research. The use of more source-like pronunciations was also observed to correlate with prescriptivism and mobility, suggesting it to carry prestige and linguistic security (Boberg 1997; 1999). It was also observed to correlate with liberal political identity (Hall-Lew et al. 2010) and with positive attitude toward the particular source of the loanword (Weinreich 1968, 27; Thomason 2001, 73; Lev-Ari and Peperkamp 2014). Previous research has also observed this variation to intuitively correlate with the speaker's degree of familiarity or fluency with the source language (Poplack and Sankoff 1984; Poplack et al. 1988; San Giacomo and Peperkamp 2008; Friesner 2009); but, it has even been observed to correlate with one's self-reporting as multilingual in general, regardless of whether that multilingualism includes the source language of the loanword of interest (Silva et al. 2011). The current study examines the way people think about this variation and whether their perception reflects the way it has been observed to pattern in production. In turn, it is examined whether these thoughts are reflected in how listeners perceive a speaker depending on whether they use more or less source-like loanword pronunciations.

There is reason to believe that the ways in which loanword variation patterns socially influence how people think about this linguistic variable and ascribe social meaning to it. In

fact, much of the motivation for the social factors considered to predict variation in Chapter 2 was the consideration of how this variation is discussed in journalism, by public figures, and in online forums.

Recall the way this variation was discussed by commenters in the online discussion forum AboveTopSecret.com (2009). Suggesting a perception of source-directed attitude, commenters indexed the use of more source-like pronunciations by American politicians as aligning more strongly with the foreign source than with the US. Reflecting a perception of globalism, it was interpreted as the speaker trying to be “perceived as cultured” or “kissing foreign butt”. It was also evaluated as more standard and carrying more linguistic security, both in positive ways (“correct”) and negative ways (“pretentious”).

Similar indexations and evaluations are observed in journalism. The headlines of two articles discussing President Barack Obama’s pronunciations of foreign words and names are displayed in Figure 13, presented along with respective descriptors used within each to characterize his apparent preference for the use of more source-like pronunciations. Each suggest similar indexations of prestige, global orientation, and linguistic security, with seemingly more positive (lefthand) and negative (righthand) stance-takings.



- 
- correct
  - a simple way of showing respect
  - baseline diplomacy
  - thoughtful



- 
- un-American
  - choice of fancy-pants internationals
  - “proper” [including scare quotes]
  - a bit more foreign-sounding

Figure 13: Journalistic reports on Obama’s pronunciations of foreign words and names. News article headlines with respective descriptors quoted from each regarding President Obama’s use of more source-like pronunciations of foreign words and names: lefthand (Lee 2009), righthand (Levenson 2014).

This study is intended to more rigorously and empirically examine how listeners per-

ceptually index loanword variation, with an aim similar to that of Chapter 2: examining the diverse potential meanings this may carry and how they might relate to each other. Alongside what might be considered *indexations* like globalism and political identity, this study will also examine listeners' *evaluations*, such as the different sentiments and framings of the commentaries mentioned above. This is in line with Preston's (1999) approach to perceptual dialectology, considering the evaluations of 'correctness', generally representing linguistic security and overt prestige, and 'pleasantness', generally representing likability and what might in certain settings be considered covert prestige. As mentioned above, the use of more source-like pronunciations is described as "correct" and even "thoughtful", suggesting a combined evaluation of correctness and pleasantness; conversely, it can also be described as "pretentious", suggesting a similar evaluation of correctness but combined instead with an unpleasant evaluation. These evaluations can also combine differently with the indexation of globalism. For example, the characterization of more source-like pronunciations as "baseline diplomacy" and "a simple way of showing respect" suggests a pleasant evaluation, while the descriptions of it as the "choice of fancy-pants internationals" suggests an unpleasant evaluation while still coupled with a globalist indexation.

One straightforward reason for this kind of study is that we know the perceptual indexation of language can not only carry and reflect social meaning but have a tangible impact. For example, in a seminal study by Purnell, Idsardi, and Baugh (1999), when placing a phone call to request to view an apartment advertised for rent, the variety of English a speaker uses significantly influences the likelihood that they will be invited to view the apartment. When using a Mainstream US English (MUSE) guise, the caller was more likely to receive an invitation in general, with a rate that remained relatively stable across neighborhoods, as compared to when the caller used a Chicano English or African American English guise. For the non-MUSE guises, the demographic makeup of the neighborhood played a role, where the caller was significantly less likely to receive an invitation to view an apartment located in a more majority-white neighborhood. Other studies observe similar

differences in treatment in the housing and home insurance markets based on a speaker's language variety (Massey and Lundy 2001; Fischer and Massey 2004; Squires and Chadwick 2006). Therefore, how a speaker's variety of speech is indexed and evaluated can have an impact on how their audience might treat them.

But, this kind of study is also of interest to sociolinguistic theory, helping us further understand how perception and production may or may not line up with each other. In many cases, these appear to line up. The use of the [m] variant of the [m]~[ɱ] variable in English is more frequent in Southern English and perceptually indexed accordingly (Campbell-Kibler 2007). African American English speakers in Philadelphia are observed not to participate in the [aʊ]→[eʊ] diachronic change in progress there, and listeners are indeed less likely to think of [eʊ] utterances as coming from an African American speaker (Graff et al. 1986). Multiple New England varieties of English are non-rhotic and, when asked to identify a non-rhotic pronunciation with a certain region of the US, listeners are more likely to say the speaker is from New England (Clopper and Pisoni 2004).

In some cases, however, these may not so neatly line up. In a benchmark study, Niedzielski (1999) finds that listeners' perceptual indexation of a linguistic variable differs from how it observably patterns in production. The use of the raised [ʌʊ] variant of the /aʊ/ diphthong is a well-known feature of Canadian English but also present in the speech of those in neighboring Detroit, Michigan. When listening to a Detroit speaker using the [ʌʊ] variant, listeners were more likely to say the speaker's realization was more [aʊ]-like than [ʌʊ]-like when told the speaker was from Detroit than when told the speaker was from Canada. Niedzielski suggests that this results from listeners' stronger associations of [ʌʊ] with Canada and that perception is therefore mediated by sociolinguistic indexation.

Similarly, Hay et al. (2006) observe that social factors influence New Zealanders' perceptual discrimination of the NEAR and SQUARE sound classes. Given that these sound classes are in the process of undergoing a merger, older speakers and those of higher socioeconomic status are more likely to maintain this distinction than younger speakers and

those of lower socioeconomic status. Hay et al. find that listeners do better at perceptually discriminating non-merged utterances of the same audio stimuli when they are simultaneously presented with pictures of speakers visually indexable as older and/or more upper-class than with speakers visually indexable as younger and/or more lower-class. This is even the case for listeners who, themselves, maintain the distinction in their own speech production. However, of course, one's own linguistic system influences this, too, with those who do not produce a distinction also not discriminating well between the sound categories. Therefore, social factors, contextual assumptions, and one's own speech variety can mediate perception.

In addition to perceptual *sound* identification and discrimination, studies also suggest that social factors can mediate listeners' *speech variety* identification. Kerswill and Williams (2002) observe that, amidst a backdrop of dialect leveling in the UK, speakers from more leveled areas do worse at identifying different dialects. Baker et al. (2009) observe that speakers from Utah are better at distinguishing Utah English from other neighboring varieties and that they rely on less stereotypical features in doing so. This suggests that linguistic variables can serve as 'in-group' markers signaling a particular identity especially well to those who share that identity while, to out-group members, the use of a particular linguistic variant may not carry as much (or at least the same) social meaning.

Besides an effect on speech variety *identification*, differences in *indexation* are also observable across different speaker groups. Yuasa (2010) examines how listeners' descriptions of a young female American English speaker are affected by whether or not she uses creaky voice. Focusing on the perceptions of this variable by college-age American listeners, Yuasa finds that creaky voice tends to be associated with education and upward mobility. This is contrary to the frequent derision of creaky voice in the public sphere and the results of Anderson et al.'s (2014) broader survey finding a speaker with creaky voice to be judged negatively, as well as less educated and less competent. Yuasa's findings suggest that creaky voice is indexed and evaluated differently by the sampled population of college-age Americans. (Anderson et al. also observe an interaction with age, where younger listeners evaluate



creaky voice more positively than older listeners.)

Therefore, how a variable is indexed and evaluated seems to depend on who is perceiving it, with both the listener's own speech variety and social identity apparently playing a role. For the study at hand, such influences will be considered: Analysis will examine not only how loanword variation is perceptually indexed and evaluated and how this lines up with the way it appears to pattern; it will also examine how a participant's own identity (in this case, political) and usage (in this case, more vs. less source-like pronunciation) mediate this perception.

The current study will also test whether the method of eliciting indexations and evaluations influences the results. In particular, this study asks: Do people think differently about loanword variation when thinking about it explicitly than when encountering it passively? First, a matched-guise study (e.g., Lambert et al. 1960; Zahn and Hopper 1985; Purnell et al. 1999; Campbell-Kibler 2007; Yuasa 2010) is used, examining whether participants' ratings of a speaker along different indexical and evaluative dimensions is influenced by whether that speaker uses more vs. less source-like pronunciations of foreign placenames. Only after this are participants made overtly aware of the variable, following up the matched-guise study with an explicit, metalinguistic questionnaire (e.g., Preston 1989; Preston 1999; Dailey-O'Cain 2000; Alfaraz 2002) directly asking participants how they might rate a speaker along different dimensions depending on whether the speaker uses more vs. less source-like pronunciations.

Multiple studies have observed similar results in the indexation and evaluation of different language varieties across methods in which participants are responding to a voice or explicitly discussing a variety of speech (Giles 1970; Coupland et al. 1994; Williams et al. 1999). On the other hand, Dailey-O'Cain (2000) observes different results between the two methods in the same study regarding the use of *like* as a focuser (e.g., "Man, get in that car, *like* now." [Underhill 1988, 239]) or a quotative (e.g., "I'm *like*, 'I know this stuff. I got a 77 last time.'" [Ferrara and Bell 1995, 266]). Matched-guise results suggest that a speaker's

use of *like* increases participants' percept of likability and solidarity with that speaker, while also perceiving the speaker as less educated; however, when asked to comment explicitly on this usage, Dailey-O'Cain notes that participants almost exclusively evaluate the use of *like* as negative, suggesting that some judgments may be stronger or, conversely, overridden when speakers are prompted to explicitly comment on language variation. Dailey-O'Cain also notes an asymmetry between explicit commentary and actual usage: In usage, younger speakers are observed to use these forms of *like* more frequently than older speakers, though with no significant difference between young women and young men; however, in explicit commentary, participants quite consistently say that it is especially young women who use them more.

The current study therefore tests if participants highlight different indexations and evaluations when commenting metalinguistically on loanword variation than when reacting to having heard it passively. Like Dailey-O'Cain suggests, some indexations or evaluations may be stronger or weaker during explicit commentary. And, any such differences may reflect the layered indexicality of interest throughout this dissertation regarding loanword variation. Indexations activated more strongly during initial exposure may be those that are more primary, like that of globalist/nationalist alignment identified in Chapter 2. Indexations showing an increased activation during metalinguistic commentary may be those that are more secondary while nonetheless socially connected and prominent, like political identity appears to be.

## **3.2 Methods**

This was an online study, with 400 participants recruited via Amazon's Mechanical Turk. Participation took under 10 minutes and participants were remunerated with a small sum of money upon completion. Participants were eligible if they were 18 years of age or older, native monolingual speakers of American English, and never previously diagnosed with a speech or hearing disorder.

Participants were told that they would listen to 2 audio clips of news reporters speaking, provide reactions about each report and reporter after having listened, and then answer some follow-up questions about themselves. Participants were instructed to be in a quiet area and use headphones for the task. (While this could not be officially verified, an additional step was taken to further encourage participants to comply: The tasks of interest were preceded by a page asking participants to identify the type and brand of headphone they were using.)

The first task was a matched-guise test with 2 trials. The participant would hear a short news report, be asked questions about what they heard, and then repeat this procedure a second time. Each news report was about one minute long. A separate speaker was used for each, but both were controlled for gender (female), age (mid 20s), and voice (aiming for a similar pitch, pitch range, voice quality, and speech rate). Both were native speakers of English, with variety of English also controlled for by instructing each to aim for a Mainstream US English “reporter” sound, taking note during recording of any potentially noticeable regional markers and training the speaker away from using them.

The content of each news report regarded international commerce. One regarded the purchase of spaceflight vehicles by some countries, considering bids from manufacturers headquartered in different countries; the other regarded an international airline expanding to include flights reaching new airports around the world. (The full text of each report is provided in Appendix D.) This allowed for the inclusion of multiple placenames in each report. The second report, the target trial of interest, contained multiple placenames that are variable in American English between more and less source-like pronunciations. These are listed in Table 15, with the locus/loci of variation between each pronunciation underlined. Two versions of this report were created: one in which all placenames were pronounced with their more source-like variants and another in which all were pronounced with their less source-like variants. The speaker read through the report multiple times, in some takes using all variants considered more source-like and in others using all considered less source-

like. One base recording was identified and Praat was used to create a second version by replacing each placename utterance with its counterpart pronunciation, with attention paid to ensuring natural transition and prosodic consistency. Half of participants heard one version, and half heard the other. The first report, that preceding the target trial of interest, also used a diverse array of placenames but those which are not considered commonly variable in American English (e.g., *Germany, Japan, The Netherlands*); therefore, participants all heard the same version of this report.

WORD	MORE ~ LESS SOURCE-LIKE	WORD	MORE ~ LESS SOURCE-LIKE
Budapest	[ˈbudəpɛst] ~ [ˈbudəpɛst]	Paraguay	[ˈpɑːrəɡwɑɪ] ~ [ˈpæːrəɡwe]
Chile	[ˈtʃɪle] ~ [ˈtʃɪli]	Quebec	[kɛˈbɛk] ~ [kwəˈbɛk]
Colombia	[kɒˈlɒmbiə] ~ [kəˈlɒmbiə]	Shanghai	[ˈʃɑŋhaɪ] ~ [ˈʃæŋhaɪ]
Iraq	[ˈrɪɑk] ~ [aɪˈræk]	Tanzania	[ˈtɑnzəˈniə] ~ [tænzəˈniə]
Pakistan	[ˈpɑkɪstɑn] ~ [ˈpækɪstæn]	Tokyo	[ˈtɒkjo] ~ [ˈtɒkio]

Table 15: Variable placenames of interest

These are the words of interest that were present in the target exposure stimulus and manipulated between more and less source-like variants. The parts underlined are those changed by the speaker/reporter when reading the two different versions.

The listener would see a page with a play button for the audio stimulus, with instructions directing them to play the report after the page finished loading but to only play it once (while allowing for them to press ‘play’ again if this was not successful the first time). After listening to each report, the participant proceeded to a page with follow-up questions. The first was a comprehension-check question, simple enough that the participant should easily be able to answer if they had listened the report but dependent enough on the content that they would likely get it wrong if they hadn’t. Any participant who failed this comprehension check regarding the target stimulus of interest was excluded from the final dataset analyzed below (§3.3.1), conservatively assuming that they may not have listened to the report at all.

After this, participants were asked for a series of reactions to the report they had listened to. These were all 7-point Likert scales, as shown in Table 16, presented in the same order (the order given in the table) to all participants. Some were designed to measure Preston’s (1999) identified ‘pleasantness’ (pleasantness, trustworthiness) and ‘correctness’

(experience, intelligence, comprehensibility) aspects of sociolinguistic evaluation. And some were designed to measure relevant social indexations: global orientation (audience span, knowledge of international current events), political orientation, and multilingualism.

Reporter:

---

How pleasant was the reporter to listen to?						
very pleasant						very unpleasant
1	2	3	4	5	6	7

---

How experienced did the reporter sound?						
very inexperienced						very experienced
1	2	3	4	5	6	7

---

How smart did the reporter sound?						
very intelligent						very unintelligent
1	2	3	4	5	6	7

---

How knowledgeable did the reporter seem regarding international current events?						
very knowledgeable						very unknowledgeable
1	2	3	4	5	6	7

---

How likely do you think it is that the reporter is multilingual (i.e., also speaks a language other than English) vs. monolingual (speaks only English)?						
surely monolingual						surely multilingual
1	2	3	4	5	6	7

Station:

---

What size station do you think this report might have been broadcast from?						
small/local station			regional/national channel		global/international network	
1	2	3	4	5	6	7

---

What political leaning do you think this station might have?						
very liberal						very conservative
1	2	3	4	5	6	7

Report:

---

How well could you understand the report?						
impossible to understand						easy to understand
1	2	3	4	5	6	7

---

How trustworthy do you think the report was?						
very reliable						very unreliable
1	2	3	4	5	6	7

---

Table 16: Report reactions

7-point Likert scales eliciting participants' reactions to the news report. All information listed was visible to the participant: aspect ("Reporter", "Station", "Report"), question, number scale, description of endpoints (and, in one case, midpoint).

After the matched-guise block, an interim question elicited participants' self-reported pronunciation of *Iraq*:

When you say the word 'Iraq', which does your pronunciation of the underlined vowel sound more similar to?

'rack'      'rock'  
○                    ○

This would allow for the analysis of how a listener's own pronunciation might influence how they perceptually index and evaluate this variable.

Then, participants were asked explicitly for their indexations and evaluations of this variable. Half of participants (within each half already grouped by which version of the report they heard) were asked about the use of more source-like variants, and the other half were asked about the use of less source-like variants. This task, like the above, elicited responses in the form of a 7-point Likert scale. Prompts and Likert scales are provided in Table 17. The first scales listed in the table again regard pleasantness (friendliness) and correctness (intelligence, education, social class) as well as a "humble ~ pretentious" scale that might be considered a combination of the two, with a "pretentious" rating suggesting an evaluation as both standard/"correct" and unpleasant. The final three listed regard the indexations of political identity, globalism/openness, and multilingualism. The order in which these ratings were displayed on the screen was newly randomized per participant.

Prompts: If you heard someone speaking English pronounce ‘Iraq’, ‘Quebec’, and ‘Chile’, for example, as /eye-rack/, /kwuh-beck/ and /chill-ee/ instead of /ear-rock/, /keh-beck/ and /chee-lay/, how might you think about them along the following factors?

If you heard someone speaking English pronounce ‘Iraq’, ‘Quebec’, and ‘Chile’, for example, as /ear-rock/, /keh-beck/ and /chee-lay/ instead of /eye-rack/, /kwuh-beck/ and /chill-ee/, how might you think about them along the following factors?

Ratings:

unkind						friendly
1	2	3	4	5	6	7
intelligent						unintelligent
1	2	3	4	5	6	7
educated						uneducated
1	2	3	4	5	6	7
lower-class upbringing			middle-class upbringing			upper-class upbringing
1	2	3	4	5	6	7
humble						pretentious
1	2	3	4	5	6	7
politically liberal						politically conservative
1	2	3	4	5	6	7
narrow-minded						open-minded
1	2	3	4	5	6	7
likely monolingual						likely multilingual
1	2	3	4	5	6	7

Table 17: Explicit Likert ratings

Prompts and 7-point Likert scales explicitly eliciting participants’ evaluations and indexations of the use of more or less source-like loanword pronunciations.

A final page elicited participants’ own political identities on a 7-point “liberal ~ conservative” Likert scale to examine whether one’s own political identity might influence how they evaluate and index this variable, especially along the factor of political identity.



### 3.3 Results

#### 3.3.1 Matched-guise

Results from the matched-guise part of the experiment suggest that the use of more vs. less source-like pronunciations of loanwords can influence how one is perceived by an audience. Figure 14 provides a violin plot of the responses to each Likert scale presented after listening to the target news report, grouped by whether participants heard the version in which the reporter consistently uses more source-like or less source-like pronunciations of the placenames within the report. A Wilcoxon-Mann-Whitney test was performed via the `wilcox.test()` function in R (as recommended by Drager [2018, 154]) to assess whether any ratings were significantly affected by the exposure condition. Table 18 provides descriptive statistics along with inferential statistics regarding each scale across exposure conditions. Three ratings show a significant effect. The strongest effect is that the reporter using more source-like pronunciations is rated as significantly more likely multilingual than when using less source-like pronunciations. The reporter using more source-like pronunciations is also considered significantly more likely to be broadcasting from a more global, wider-spanning station and to have more reporting experience.

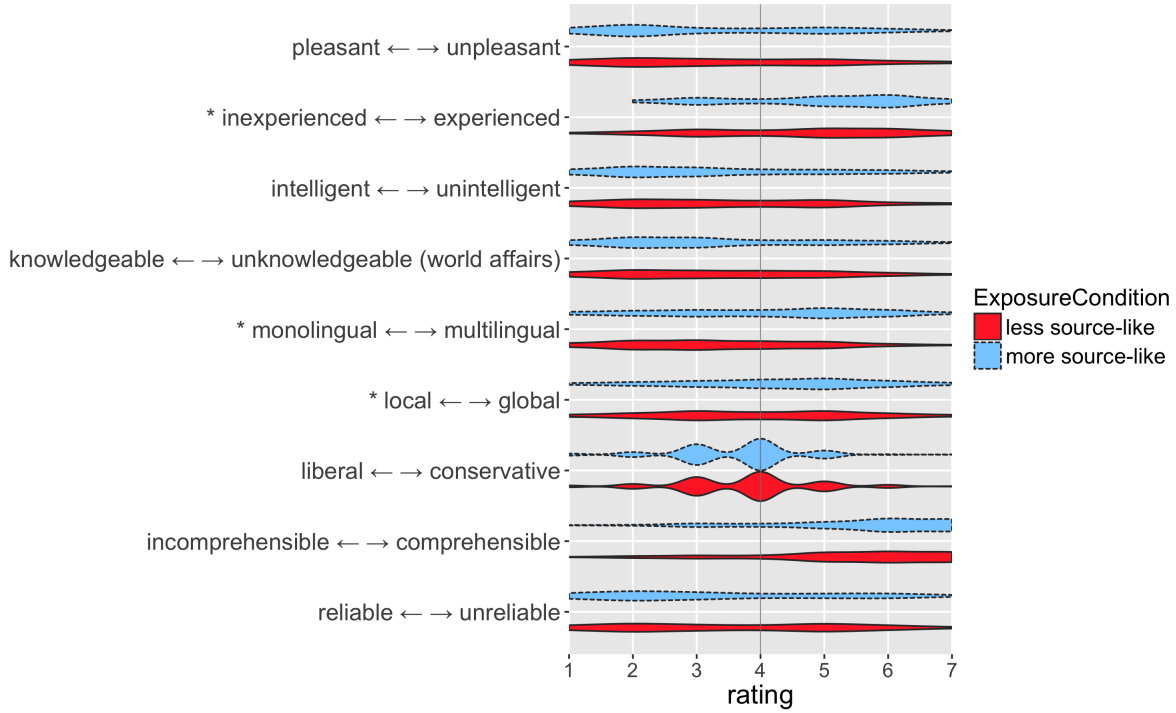


Figure 14: Matched-guise results

Likert ratings of the reporter, grouped by whether one heard the reporter use more source-like or less source-like pronunciations of placenames during the report. Ratings for which results were significantly influenced by exposure condition are marked with <\*>.

SCALE	MORE SOURCE-LIKE		LESS SOURCE-LIKE		DIFFERENCE		
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\Delta\mu$	$W$	$p$
pleasant ~ unpleasant	3.224	1.712	3.411	1.711	0.187	21545	.236
* inexperienced ~ experienced	4.917	1.417	4.599	1.574	0.318	17911	.045
intelligent ~ unintelligent	3.234	1.649	3.416	1.622	0.182	21619	.213
knowledgeable ~ unknowledgeable (re: world affairs)	3.229	1.581	3.386	1.579	0.157	21459	.268
* monolingual ~ multilingual	4.083	1.626	3.34	1.461	0.743	14747	2.006e-6
local ~ global	4.302	1.507	3.949	1.544	0.353	17471	.017
liberal ~ conservative	3.615	0.956	3.761	1.054	0.147	21646	.184
incomprehensible ~ comprehensible	5.605	1.416	5.391	1.493	0.214	18404	.113
reliable ~ unreliable	3.366	1.798	3.584	1.755	0.218	21618	.214

Table 18: Matched-guise statistics

One notable aspect of these results is that even the significant effects observed are not large. They are quite subtle, suggesting that the variable itself was not very salient to listeners and/or that its percept is not polarizing and leading to drastic, uniform effects.

Turning to the explicit, metalinguistic results below will further contribute to these considerations (§3.3.2): Maybe people do have stronger opinions about what this variation means when they think about it explicitly. A further consideration is that this variable may be more noticeable to some listeners than to others and/or that some listeners' ratings might be influenced by the variant they heard while others might be more indifferent. This consideration will be given due attention when analyzing how participants' own political identities and usage interact with these effects (§3.3.3).

Another notable aspect is that the political rating was not affected. In fact, the results of this rating appear to be quite strongly clustered at the center. It is possible, though, that participants simply did not feel comfortable opining about politics, given its charged and therefore sensitive nature. The other possibility, however, is that these indexations of multilingualism, global orientation, and reporting experience are stronger and more consistent than any political indexation of this variable, much like what was observed in Chapter 2 regarding how this variation patterns in production.

These significant effects observed are in line with previous findings regarding how this variation patterns in production. Silva et al. (2011) observe that speakers who self-report as multilingual are those who are more likely to use the more source-like pronunciations of *Iran* and *Iraq*. As observed in Chapter 2, those who are more globalist-aligning are those who are more likely to pronounce a broad range of established loanwords with their more source-like variants. And Boberg (1997; 1999) suggests that there is more linguistic security amongst American English speakers in the use of [ɑ], the positedly more source-like pronunciation for the 'foreign (a)' found in loanwords. This linguistic security seems reflected in the judgment of a user of more source-like pronunciations as having more experience in reporting, a job usually requiring one to speak "well", "correctly", or "articulately" (however these classifications are construed by societal standards). Furthermore, the nexus of these indexicalities suggests that this variation may have an influence on one's capital in what Bourdieu (1977) terms the 'linguistic market', a setting in which one speaking a certain way

could prove either costly or opportunely to them. In particular, it appears that the use of more source-like pronunciations gains one capital in what might be considered the *global* linguistic market (Piller 2001; Zhang 2005; Blommaert 2010), leading one to be perceived as more linguistically flexible and accommodating to a wider, more linguistically and/or culturally diverse audience. This idea will be further discussed below (§3.4).

### 3.3.2 Metalinguistic subjective evaluation

Figure 15 provides a violin plot of the responses to each Likert scale in the explicit, metalinguistic ratings task. In this case, all indexations tested appear to be activated, and many with stronger effects. All exhibit a significant effect in the expected direction, with descriptive and inferential statistics provided in Table 19.

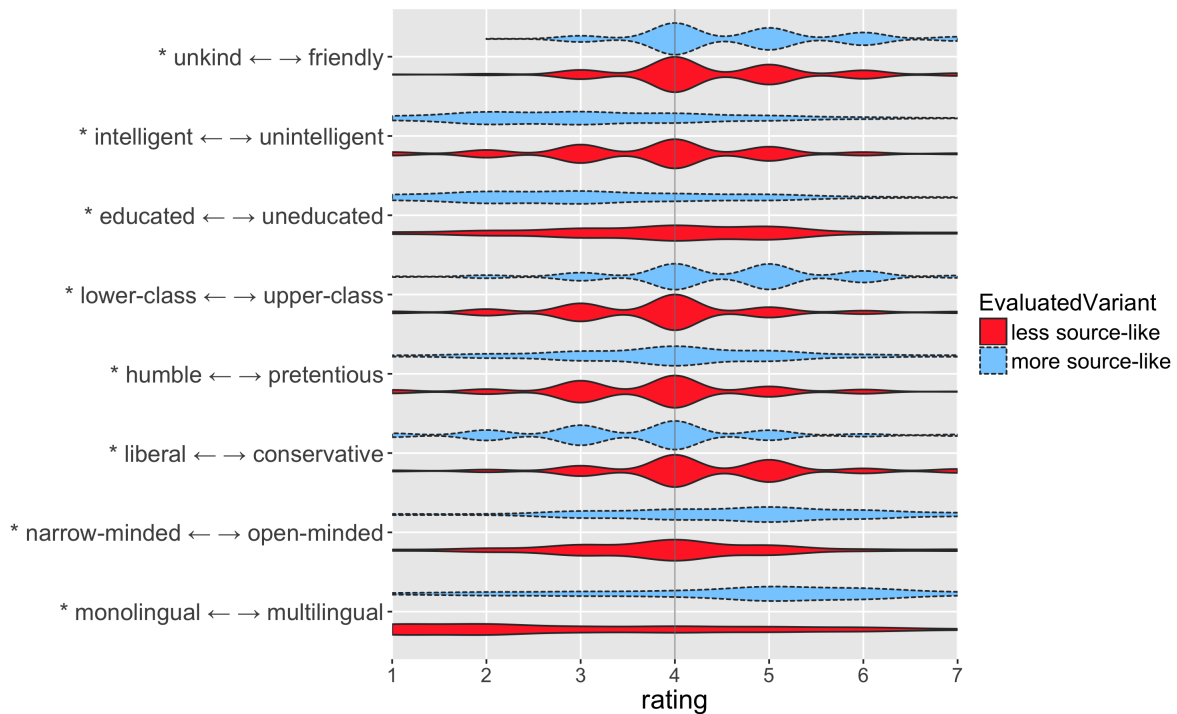


Figure 15: Explicit metalinguistic rating results

Likert ratings of general loanword variation, grouped by whether one was prompted to rate the use of more vs. less source-like pronunciations. Ratings for which results were significantly influenced by variant evaluated are marked with <\*>.

SCALE	MORE		LESS		DIFFERENCE		
	SOURCE-LIKE		SOURCE-LIKE		$\Delta\mu$	$W$	$p$
	$\mu$	$\sigma$	$\mu$	$\sigma$			
unkind $\sim$ friendly	4.69	1.034	4.406	1.043	0.284	17341	.0095
intelligent $\sim$ unintelligent	3.135	1.336	3.748	1.214	0.613	25933	4.36 $e$ -7
educated $\sim$ uneducated	3.080	1.398	3.911	1.274	0.831	27307	4.52 $e$ -10
lower-class $\sim$ upper-class	4.505	1.143	3.807	1.119	0.698	13005	1.15 $e$ -10
humble $\sim$ pretentious	3.985	1.217	3.713	1.123	0.272	17547	.0172
liberal $\sim$ conservative	3.580	1.237	4.317	1.141	0.737	27155	5.01 $e$ -10
narrow-minded $\sim$ open-minded	4.655	1.332	3.941	1.183	0.714	13519	3.43 $e$ -9
monolingual $\sim$ multilingual	4.695	1.508	3.019	1.765	1.676	9798	1.34 $e$ -19

Table 19: Explicit metalinguistic rating statistics

The use of more source-like pronunciations is rated as significantly more intelligent, educated, and upper-class, which once again suggest a greater degree of linguistic security, or ‘correctness’ (à la Preston 1999). It is also rated as more friendly, which could suggest a greater degree of ascribed ‘pleasantness’ (à la Preston 1999); however, it is also rated as significantly more “pretentious”, suggesting that while more source-like pronunciations might be considered more “correct” in general, this may be combined with a negative connotation and not considered completely/consistently “pleasant”. Similar to the matched-guise results, one using more source-like pronunciations is also rated as significantly more open-minded (a rating intended to access a globalist-aligning indexicality) and more likely multilingual.

Unlike the matched-guise results, political indexation now exhibits a significant effect, with the use of more source-like pronunciations rated as more politically liberal. This helps address a question that arose when analyzing the matched-guise results: whether participants are uncomfortable opining about politics. Here, it seems that they are comfortable opining about politics, which in these metalinguistic results shows one of the stronger effect sizes ( $\Delta\mu=0.737$ ). The difference between these results and the matched-guise results suggests that loanword variation does carry a political indexation, but that this is only (or at least more strongly) activated when one thinks overtly about it. This might suggest a perceptual kind of indexical order (Silverstein 2003), parallel to that observed in production, which will be further discussed below (§3.4).

### 3.3.3 Considering factors of the participant

Further analysis considers how a participant’s ratings might be different depending on their own political identity and their own pronunciation of  $Ir[a \sim \text{æ}]q$ . Results (Figure 16) suggest that the full 7-point spectrum of political identity was represented across participants while recognizably skewed, with more participants identifying as liberal than as conservative.<sup>4</sup> Results also suggest that within each political identity group there was at least some observable variation, with at least a few per group self-reportedly using the [ɑ] pronunciation of  $Ir\underline{a}q$ . Like the results of Chapter 2 (§2.3.3.2), though, across all groups the majority reports to use the [æ] pronunciation.

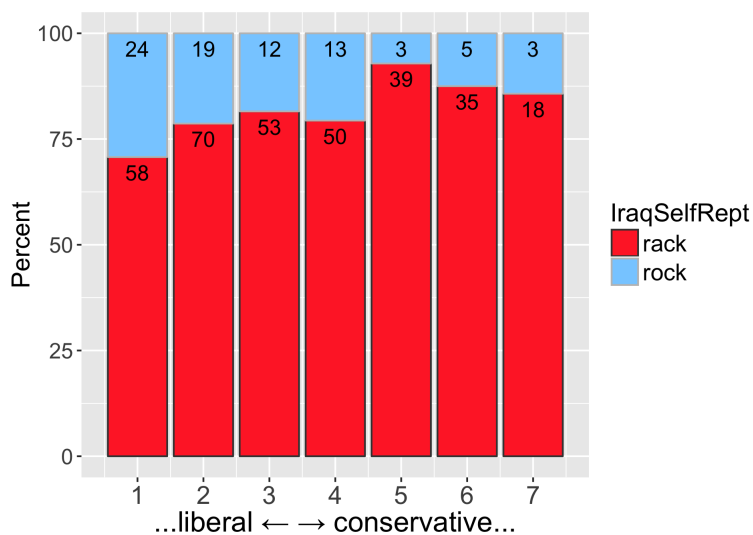


Figure 16: Distribution of participants: political identity  $\times$   $Ir[a \sim \text{æ}]q$  pronunciation  
 Distribution of participants across the factors of self-reported political identity and  $Ir[a \sim \text{æ}]q$  pronunciation. Stacked bar height represents percentage of participants within each political identity group. Each bar is labeled with the number of participants per political identity  $\times$  pronunciation group.

In the following analysis, linear regression models, using the `lm()` function in R, were generated for the ratings from both tasks. For each model, an interaction term between each factor of the participant and the exposure condition was included to test if any effects were especially strong for participants of a certain identity or self-reported pronunciation. First

<sup>4</sup>This is not surprising, seeing that MTurk participant pools consistently skew liberal, while not significantly differing in behavior from other non-MTurk studies examining political identity and behavior (Levay et al. 2016).

(§3.3.3.1), the results of the explicit metalinguistic rating task are analyzed and discussed to provide a sense of how different participants might think differently about this variation. Then (§3.3.3.2), the results from the initial matched-guise experiment are inspected for similar effects to test if they extend to participants’ processing of this variable without thinking about it explicitly. Only significant or near-significant results will be discussed below. Full outputs of statistical models are provided in Appendix E.

### 3.3.3.1 Metalinguistic ratings by factors of the participant

Two ratings are identified to exhibit a significant interaction with the participant’s own pronunciation of *Ir*[a ~ æ]*q*. For participants who, themselves, use the more source-like [a] pronunciation of *Ir*a*q*, the rating of a speaker using less source-like pronunciations as less intelligent is especially strong ( $\beta=0.619$ ,  $\sigma=0.215$ ,  $t=2.88$ ,  $p=.0042$ ). This effect further supports the notion that the use of more source-like variants carries greater linguistic security: Those who use the more source-like pronunciation more strongly judge a speaker using the other as “unintelligent”, which Preston (1999) identifies as a proxy judgment for the general ‘correctness’ evaluation. This also extends to the more specific indexation resembling that of globalist/nationalist alignment, with those who use the more source-like pronunciation judging those who use the less source-like as less open-minded ( $\beta=0.471$ ,  $\sigma=0.212$ ,  $t=2.218$ ,  $p=.027$ ). This may be especially relevant when considering [a]-pronouncers who simultaneously identify as politically conservative: One identifies this as the more ‘correct’ and/or more global/open-minded pronunciation even if they are politically conservative. And, given Chapter 2’s findings that globalist/nationalist alignment is the strongest predictor of this variation, it appears that those who are more globalist-aligning (whether simultaneously liberal-identifying or not) consider the use of more source-like loanword pronunciations a way to signal such alignment.

The political identity of the participant is observed to significantly interact with the results of a rating scale. Participants who identify as more liberal more strongly rate the

use of less source-like pronunciations as coming from a speaker with a lower-class socioeconomic background ( $\beta=0.125$ ,  $\sigma=0.044$ ,  $t=2.857$ ,  $p=.0045$ ). This could be interpreted as another reflection of linguistic security, where liberals' stronger association of less source-like pronunciations with lower socioeconomic status is a proxy for liberals more strongly considering such pronunciations as less 'correct'. Or, this could more directly mean that liberals are more attuned to how socioeconomic status and/or mobility might influence this variation, where the use of more source-like pronunciations recognizably carries overt prestige and those from a more upper-class background might have more access to this standard. Of possible relevance to these considerations is the absence of such an interaction with the "intelligence" rating. While those who use the more source-like pronunciation of *Iraq* judge the use of less source-like pronunciations as less intelligent (as noted above), this judgment is not stronger amongst those who are more liberal-identifying. This suggests that [a]-pronouncers of *Iraq* across all political identities (29% of the most liberal-identifying, 14% of the most conservative-identifying) have greater linguistic security in their pronunciation but that this is not necessarily greater still for those who are more liberal-identifying. This may therefore discourage us from interpreting the interaction of the social class rating with political identity as a similar proxy for 'correctness'.

Finally, both participants' own political identities and their own pronunciations of *Ir[a ~ æ]q* are identified to significantly influence how they index this linguistic variable with political identity. More liberal-identifying participants more strongly rate the use of more source-like pronunciations as coming from a more liberal-identifying speaker ( $\beta=0.172$ ,  $\sigma=0.046$ ,  $t=3.717$ ,  $p=.00023$ ), as well as those who use the more source-like [a] pronunciation of *Iraq* themselves ( $\beta=0.416$ ,  $\sigma=0.198$ ,  $t=2.106$ ,  $p=.036$ ). This suggests that the political indexation of this variable is stronger for liberals, considering the use of more source-like pronunciations as an in-group marker by those who use such pronunciations themselves. Conversely, though, more conservative-identifying participants less strongly rate the use of more source-like pronunciations as coming from a more liberal-identifying speaker. As



considered above regarding the “narrow-minded ~ open-minded” scale, this could suggest that more globalist-aligning conservatives consider this more a marker of such alignment than of political identity more generally.

### 3.3.3.2 Matched-guise ratings by factors of the participant

The same analysis was applied to the rating results of the matched-guise task. For the strongest effect observed above, that regarding the reporter’s likelihood of being multilingual, no significant interaction is identified. However, both factors of the participant are identified as significant influences along the “local ~ global” rating: The rating of a reporter using less source-like pronunciations as reporting from a more local and less global broadcasting station is especially strong for listeners who identify as more politically liberal ( $\beta=0.164$ ,  $\sigma=0.059$ ,  $t=2.777$ ,  $p=.0058$ ), as well as for those who use the [ɑ] pronunciation of *Iraq* rather than the [æ] pronunciation ( $\beta=0.643$ ,  $\sigma=0.299$ ,  $t=2.148$ ,  $p=.032$ ).

A similar significant effect is observable for the “inexperienced ~ experienced” rating: The rating of a reporter using less source-like pronunciations as having less reporting experience is especially strong for listeners who identify as more politically liberal ( $\beta=0.141$ ,  $\sigma=0.058$ ,  $t=2.422$ ,  $p=.016$ ), and a near-significant trend suggests that this may also be the case for those who use the [ɑ] pronunciation of *Iraq* rather than the [æ] pronunciation ( $\beta=0.557$ ,  $\sigma=0.295$ ,  $t=1.887$ ,  $p=.059$ ). Additionally, while not significant, another trend was observed suggesting that more liberal-identifying participants may have considered a reporter using less source-like pronunciations to be reporting from a more conservative-leaning station ( $\beta=0.07$ ,  $\sigma=0.039$ ,  $t=1.78$ ,  $p=.076$ ).

The combination of these effects of participants’ political identities and self-reported pronunciations is similar to their effects observed above regarding participants’ metalinguistic judgments about loanword variation. Those who are more liberal-identifying and who, themselves, use the more source-like pronunciation of *Iraq* seem especially attuned to another’s use of more source-like pronunciations and what this might mean about the speaker.

In fact, they may be the only listeners whose evaluations of a speaker are influenced by their pronunciations of loanwords when not thinking about it explicitly: In the models of the metalinguistic ratings discussed above, the main effects were still significant in the same direction along with the interaction terms included; in the matched-guise dataset, however, only the interaction terms are identified as significant effects when included alongside the main effect of exposure condition. It therefore appears that liberal-identifying [ɑ]-pronouncers of *Iraq* are those whose perceptions are most affected by whether a reporter diverges from them and uses less source-like pronunciations, bringing them to assume that the reporter is not as experienced and reporting from a less widespanning (and possibly more conservative-leaning) station.

### 3.4 Discussion and Conclusions

The results of this study suggest that loanword variation in American English is perceptually indexed in ways similar to how it patterns in speech production, as observed in Chapter 2. The use of more source-like pronunciations is observed in Chapter 2 to pattern with globalist alignment and liberal political identity, and in this study it appears that a speaker using more source-like pronunciations is perceptually indexed as more globalist-aligning, globally oriented, and politically liberal. It is also associated with overt prestige and linguistic security, resembling the observation in Chapter 2 that those who hold a more prescriptive ideology regarding language are those likely to use more source-like pronunciations. And, with loanwords being the manifestations of language contact, it is not surprising that the use of more source-like pronunciations is also associated with multilingualism, similar to the observations made by Silva and colleagues (2011) that American English speakers self-reporting as multilingual are more likely to use the more source-like pronunciations of *Iran* and *Iraq*.

Results of this study also suggest that perception can reflect the layered nature of sociolinguistic indexicality. In Chapter 2, variation along political identity was examined

with attention to comparing political identity with other social predictors. It was observed that globalist/nationalist alignment is a stronger predictor of loanword variation in production, with much of the variation observable along political identity falling out from political identity's correlation with globalist/nationalist alignment. However, some significant variation was still observable along the factor of political identity, especially for loanwords from sources of more political charge in the US. This was interpreted to suggest that the political indexation is second-order (Silverstein 2003) to that of globalist/nationalist alignment, exhibiting loanword variation as a result of its correlation with globalist/nationalist alignment while still subject to social reconstrual (Silverstein 2003; Eckert 2008), given this correlation as well as considering the social salience of this correlation in the US political zeitgeist.

In the current study of perception, the results may be interpreted as reflecting a similar indexical order in a different way. In an implicit, matched-guise test of perceptual indexation, the use of more source-like pronunciations appears to be associated with multilingualism and global orientation. A political indexation is only observed to surface when participants are explicitly prompted to comment on what loanword pronunciation may tell them about a speaker. Therefore, this indexation does not seem to be as strongly or primarily activated upon exposure; only when thinking metalinguistically (or what we might also consider metasocially) do participants suggest a political meaning, possibly as a result of the association between US politics and globalism/nationalism being more strongly activated upon more direct, focused contemplation.

Aside from what this reveals about the layered indexicality of loanword variation specifically, these observations and interpretations may be extended more generally to the field of sociolinguistic perception and the methodology used to study it. Research using implicit methods of measuring sociolinguistic perception has gained much ground (e.g., Lambert et al. 1960; Cooper 1975; Purnell et al. 1999; see Campbell-Kibler 2010 for a review). And there is reason for being cautious about asking untrained non-linguists to consciously surmise about the patterns of language, be they social or grammatical—for making sure it

is linguists doing the language analysis. As the following excerpt from the introduction of Bauer and Trudgill's edited volume titled *Language Myths* opines, non-linguists may more likely be wrong, misinformed, or oversimplifying when providing their own interpretations or analyses of linguistic phenomena.

[I]f you want to know how language works you should ask a linguist and not someone who has used language successfully in the past. [...] As linguists, we are very much aware that ordinary people have some well-established ideas about language. We meet these ideas when non-linguists talk to us at parties, in the common rooms of universities, from members of our families and in the media. Some of these ideas are so well established that we might say they were part of our culture. It is in this sense that we refer to them as myths [...]. But in very many cases, our reactions, as professionals, to these attitudes, to these myths, is: 'Well, it's not actually as simple as that.' Sometimes we think that the established myth is downright wrong. Sometimes we think that two things are being confused. Sometimes we think that the implications of the myth have not been thought through, or that the myth is based on a false premise, or that the myth fails to take into account some important pieces of information. (Bauer and Trudgill 1998, xvi)

The results of this study motivate such a caution with respect to indexical order: When prompted to comment explicitly on a variable, participants may be adding indexations that can indeed be associable with a linguistic variable but in a second-order fashion, by transitive association with more direct, primary social indexations. However, as the above excerpt also suggests, even if the interpretations by non-linguists are not necessarily accurate, they can be meaningful and even influential. As Preston opines, "the discovery of what non-linguists believe about and do with language ('folk linguistics' in general) is an equally important issue" (1996, 72). But, as Preston also suggests, "folk linguistic responses to regional varieties [and surely varieties of other kinds] can profit from being investigated by using a variety of research methods" (1999, 131). The current study's multi-method approach shows the importance of being aware that metalinguistic commentary may reveal a broader array of indexations associated with a variable, but it may miss the nuanced relationship of those indexations and how some may be more primary than others. The combined results of this study suggest that metalinguistic commentary can be complemented by methods designed to avoid transparency of the variable, isolating the indexations and evaluations of a variable that are activated more directly when encountered in actual usage. And, returning to the

specific linguistic variable at hand, the fact that participants do explicitly index loanword variation with political identity could be a foreshadowing of the future of this variable, suggesting that its political indexation could gain strength and prominence.

Aside from the method of elicitation, factors of the individual appear to influence the indexation of this variable, even within results elicited using the same method. Some indexations are activated more strongly for certain listeners, suggesting that this variation may be more or differently meaningful for some than for others. When examining the effect of a participant's own usage, it appears that the use of more source-like pronunciations carries greater linguistic security. Those who use the more source-like [ɑ] pronunciation of *Iraq* are those who consider a reporter using less source-like placename pronunciations as less experienced and, in explicit ratings, more strongly judge the use of less source-like pronunciations as less “intelligent” and more “narrow-minded”. When examining the effect of a participant's political identity, results suggest that the use of more source-like pronunciations is an in-group marker for those who identify as politically liberal. Those who identify as liberal, and especially those who are also [ɑ]-pronouncers of *Iraq*, are those who more strongly consider the use of more source-like pronunciations to signify that the speaker using them is also politically liberal. That is, liberals appear to consider this an ‘in-group’ marker, as discussed above (§3.1). However, they also appear to recognize that this may be a prestige marker, as well: that a speaker using less source-like pronunciations may have lower socioeconomic status.

A consistent effect that is not mediated by factors of the participant is the indexation of multilingualism with the use of more source-like pronunciations. This indexation, however, might not be as simple as it seems. Recall that the rating regarded the likelihood that one is “multilingual (i.e., also speaks a language other than English) vs. monolingual (speaks only English)”. When designing the prompt for this rating, this wording was expressly chosen to avoid the potential ambiguity of whether this rating regarded the likelihood of one speaking another language or whether it regarded one's degree of familiarity with any language(s)

other than English more generally. However, as also expressly chosen in the design of this study, all pronunciation variants were still readily available within the native English sound system. It is not the case that, to utter the more source-like variants, the speaker was adjusting her phonetic or phonological system to use sounds unavailable in the English inventory. Furthermore, the words chosen were sourced from a wide variety of languages, and one would likely not assume that the speaker holds fluency in all of these languages simultaneously.

Some nuance is therefore motivated when considering what this rating of multilingualism actually means to participants. Especially when we consider this rating in combination with the other indexations of global orientation, open-mindedness, reporting experience, and more general linguistic security, it appears that the use of more source-like loanword pronunciations carries capital in what might be considered the global linguistic market (Bourdieu 1977; Piller 2001; Zhang 2005; Blommaert 2010). More strongly associating a reporter using such pronunciations with multilingualism may not necessarily mean considering her fluent in a large and diverse suite of languages but, rather, considering her more able to navigate multilingual settings. Using more source-like pronunciations may lead one to be perceived as more linguistically flexible and accommodating to a wider, more linguistically and/or culturally diverse audience.

The journalistic articles mentioned above (§3.1) suggest a similar interpretation. Each considered the use of more source-like pronunciations to reflect a global orientation and alignment. One (Lee 2009) highlighted this notion that it may carry capital in the global linguistic market, with characterizations of “thoughtful”, and “baseline diplomacy”. The other, however, characterized this with a more negative evaluation as the “choice of fancy-pants internationals”. This suggests that the use of more source-like pronunciations could *cost* one linguistic capital at the more national/local level, further supported by concurrent characterizations within that same article as “un-American”.

This furthers our understanding (while raising further questions) of the global linguis-

tic market and particularly the relation of English to it. It is clearly the case that much of this market is English-dominated and continually reinforcing a system in which speaking English is a highly prized gateway to global linguistic capital (Phillipson 1992; 2009; Block et al. 2012). Piller (2001), for example, observes that Germans place a high value on German-English bilingualism, with the use of English increasing and being highly commodified in commercial advertisements but also beyond such a directly commercial setting. Zhang (2005), however, suggests that not all which can be considered global linguistic capital is English, observing the use of more “cosmopolitan”-indexed Mandarin features to be employed in more transnational business settings (rather than those more local). The observations of the study at hand may further suggest that English is not the only potential capital-carrier in the global linguistic market: that the incorporation of (or at least stronger faithfulness to) non-English forms *in* English may carry global linguistic capital.

However, some significant caveats must be taken into account when considering how the faithfulness to non-English words in an English-matrix setting may carry global linguistic capital. For one, the listeners in this study most definitely do not represent a global pool, so the results of this study cannot be inferred to suggest how the global population might react to and evaluate the variation of loanwords in English speech. The interpretation of these findings with respect to global linguistic capital should therefore be narrowed: While we may not know how people across the world evaluate this variation and may or may not ascribe it linguistic capital, we see that American English speakers *consider* the use of more source-like pronunciations of loanwords by a fellow American English speaker a way to gain global linguistic capital. But, given that they, as native English speakers, are already prime carriers of such capital, such considerations may not be trivial:

“The two aspects of linguistic authority or hegemony, then, are knowledge or control of a standard, and acknowledgment or recognition of it [... A] variety may be said to be hegemonic even if a large part of the population does not control that variety; that would, in fact, constitute the typical situation.” (Woolard 1985, 741)

Those who are in control of the standard or capital are those who get to define how capital

is gained, so long as this control is recognized. Therefore, the opinions expressed by these capital-carriers in the global linguistic market may be relevant, and possibly even influential, to how other speakers within the market consider capital to be gained.

Another caveat is the fact that the “reporter” participants in this study heard during the matched-guise portion used what would be considered Mainstream US English and, furthermore, she may likely have also been assumed to be white (à la Purnell et al. 1999), in spite of the fact that she was not. A cautionary asterisk must then be placed on the interpretation of these results, motivating a more narrowed scope: The faithfulness to non-English forms in (Mainstream US) English (as spoken by a likely white-indexed speaker) may carry global linguistic capital. Regarding global multilingualism more generally, Flores (2015) notes that not everyone benefits in the same way from English+Other bilingualism. In line with others (Valdés 1997; Petrovic 2005; Ricento 2005), Flores raises a criticism of the language-as-resource orientation regarding the promotion of bilingual education (see Ruiz 1984), suggesting that “[I]n a society with hierarchies created by hegemonic Whiteness, language as a resource for all is likely to benefit those who most closely fit the ideals of hegemonic Whiteness—namely, White middle- and upper-class students and their families” (Flores 2015, 31-32). In other words, and in relation to the current language contact phenomenon of interest, one may more likely (or exclusively) benefit from the “resources” of another language—in this case, the faithfulness to the sound forms of words borrowed from other languages—only if added on top of the majority, white-indexed language or language variety. Future research may therefore examine how the perceptual indexation and evaluation of loanword variation in English could differ when the speaker is presumed to be a Person of Color and/or a speaker of a variety other than that which might be considered Mainstream US English.

A similar extension is motivated with respect to the listener. One who is a speaker of the source language, or who otherwise personally identifies with that language or people/places associated with it, may have particular attitudes regarding how a loanword is



pronounced. As Lee (2009) reports about President Obama’s pronunciation of *Pakistan* with the more source-like [ɑ] rather than [æ], “Pakistanis have told the White House they appreciate it.” However, this may not so consistently be the case, possibly interacting with the identity of the listener and/or contextual factors. And other factors related more directly to the loanword, like the degree of establishment or to what degree one might adjust their sound system to replicate the source form, add layers of complexity to how this variation may be indexed and evaluated. This chapter therefore ends with a motivation for future studies to consider and tease apart the multiple factors at play.

When speaking about loanword variation with non-linguists, many may point to the iconic representation in a Saturday Night Live sketch (“Antonio Mendoza” 1990) in which actor Jimmy Smits portrays Antonio Mendoza, a Latino man newly hired at a TV station in which all of his white coworkers are observed to use non-established pronunciations of Spanish loanwords, inferrably trying (with great effort, while not always with the greatest accuracy) to replicate how such words might be pronounced in Spanish. Given the comedic platform that is Saturday Night Live, this is unsurprisingly done excessively and emphatically; however, the message is clear that this behavior is deemed awkward. After observably escalating frustration, Antonio finally speaks up to his coworkers about this:

“I’ve just been noticing that you guys really are up on your Spanish pronunciation. ... But—um— if you don’t mind— if you don’t mind me saying so, sometimes these Spanish words, when you take ’em and you sort of— kind of— um— over-pronounce them and it’s... kind of annoying.”

Perhaps this annoyance could be construed as only a result of how fervent and emphatic the others are about this practice of “over-pronunciation”. For example, with most every token there is some kind of anticipatory pause coupled with a vigorous movement of the head and an increase in volume by the speaker. But perhaps there is more to it. There is a very wide range of words subject to this practice: placenames not solely pertaining to those that are Spanish-language-dominant (e.g., *Los Angeles*), Latino foods, an American sports team, and others such as *Camaro* (the Chevrolet car model) and *tornado*. The words also display a range of establishment, from those very infrequently used in English to those

which non-etymologists might not even posit as being loanwords. Even a variety of contexts is present within the short sketch: It starts by showing the TV with a newscaster from the station reporting from Managua and then the rest takes place in the office of the TV station with those physically present.

The laughter of the audience and cringes by Antonio could serve as a clue to how these different factors affect the percept of this variation. The clear outlier candidates such as *Camaro* and *tornado*, with dubitable and/or at least likely unposited connection to Spanish by the audience, garner much response. The other words that seem to garner the most response are those whose meaning is associated with the US and not a Spanish-language-dominant country: proper names representing US-pertinent locales or entities such as *San Francisco*, *Los Angeles*, and *Broncos* (the Denver football team). Amongst the other words, those of higher English frequency and whose unnativized forms require a fuller phonetic/phonological code-switch to incorporate non-English sounds also appear to have a relatively stronger cringe-worthiness shared by Antonio and the audience: e.g. *burrito*, which satisfies both of these factors by being frequent and being pronounced by the speaker with a saliently non-American English rhotic trill.

Something potentially notable is the pronunciation of *Managua*, *Nicaragua* by the reporter in the opening of the sketch, which does not elicit any laughter by the audience, even though the pronunciation difference is more of a code-switch of the sound system rather than a choice between two equally established and English-compliant variants ([ma'nagwa nika'ragwa], cf. [mə'nægwə nɪkə'ɹægwa]). Granted, this could be due to its early positioning within the sketch, possibly before the audience catches on to the recurring locus of its humor. When watching, though, the reporter's unnativization seems intuitively less awkward, inappropriate, or cringe-worthy. This may be a result of the constellation of potentially influential factors lining up in this instance, leading this utterance not to be as jarringly considered an "over-pronunciation". One factor is that the words she unnativizes are not those referring to US-pertinent locales, as discussed above, and not highly frequent in English.

Another possibly influential factor is that she is reporting from Nicaragua. Her unnativization, then, may be more licensed as “doing as the Romans”, given her location. Her speech being public-oriented, and internationally so, rather than casual could also play a large part.

This final consideration ties back to the current study’s analysis. The matched-guise stimulus framing was a news report, mainly chosen for two reasons: 1) so that it would feel somewhat reasonable and natural to be asking listeners for ratings about the speaker, and 2) so that it would feel natural that many foreign placenames are being uttered, which would otherwise be rather uncommon in more everyday speech. If this were more everyday speech, though, (like that in the workplace at the TV station between coworkers while not broadcasting) some other indexations might have been more strongly activated without the elicitation of explicit metalinguistic commentary.

Future research teasing apart such potential confounds will further our understanding of how loanword variation is perceived. These include the identity and language (variety) of the speaker, the identity and language (variety) of the listener, the degree of establishment of the loanword, the degree of foreignness associated with the loanword (e.g., Spanish-sourced US placename vs. Latin American placename vs. food name), the context and audience, and to what degree the speaker adjusts their sound system away from that of English to incorporate non-English sounds or phonotactics (and the possible salience differential between certain kinds of adjustments).

While the following study (Chapter 4) does not continue this program examining listeners’ perceptual indexations and evaluations, it does rigorously account for one of these factors. The degree of establishment of the loanword is controlled by examining how speakers reproduce the source forms of new loanwords they encounter. The previous study (Chapter 2) set the stage by examining how established loanwords vary along different social dimensions. The current study examined how listeners perceive loanword variation and how such perception relates to the social patterning of this variation in actual usage, suggesting how certain indexations may be gained, as well as how they might be influential regarding the

future of this variation. The following study cycles back to consider how loanword variation may arise in the first place, testing how faithful speakers are to the sound forms of new loanwords and how this, too, may be mediated by the same social factors considered throughout the studies conducted in this dissertation so far.

## Chapter 4

### Loanword adaptation as socially mediated sound imitation

#### 4.1 Introduction

The borrowing of a loanword, by definition, entails the imitation of the sound form paired with the word's meaning in the source language. As discussed in Chapter 1 (§1.3.1), this makes loanword variation form a not-quite-arbitrary sociolinguistic variable, where different forms can be analyzed by how close a replication they are of the source form. In many cases, especially across languages, this can be explained by differences between the borrowing languages' sound systems. However, there are cases of loanword variation *within* a single language, where different pronunciations appear that can be analyzed as more or less source-like (Weinreich 1968, 27; Poplack et al. 1988; van Oostendorp 1997; Thomason 2001; Silva et al. 2011). These prompt the question as to whether or not there are social factors mediating this replication of the source form. There are multiple studies outside the realm of loanword adaptation suggesting that sociolinguistic variation more generally can be attributed to the way social factors can induce or mediate the imitation of the linguistic features of another. This will be reviewed below (§4.1.1), motivating the further extension of this kind of imitation analysis to loanword adaptation as a way to possibly explain how sociolinguistic loanword variation arises (§4.1.2). These considerations together will motivate the empirical pursuit of this study: to test if the same social factors observed in Chapter 2 to predict the variation of established loanwords between more and less source-like pronunciations similarly influence speakers' replication of the exposure forms of new loanwords they encounter.

### 4.1.1 Socially mediated phonetic imitation

A loanword is the result of the borrowing of a lexeme, including the source language’s surface form of it. The replication of this source form, of course, is constrained by the borrowing language’s sound system. However, a similar phenomenon of sound replication can be observed even within the *same* language but across *speakers*. A large body of research has found that the phonetic detail of utterances by one speaker can influence the phonetic realization by another. For example, Fowler et al. (2003) conduct a listen-and-repeat shadowing experiment, where native English-speaker participants hear [p<sup>h</sup>a], [t<sup>h</sup>a], and [k<sup>h</sup>a] utterances and are asked to repeat what they hear. In some exposure trials, the voicing onset time (VOT) after the release of the stop is manipulated to be longer than that in other trials. It is observed that participants’ own productions match what they hear, having longer VOTs when they hear stimuli with longer VOTs and vice versa, in spite of this phonetic variation not being contrastive in English. This has been observed in a variety of methodologies: further repetition tasks involving nonce sound sequences (e.g., Gentilucci and Bernardis 2007) and real words (Goldinger 1998, Namy et al. 2002), as well as interlocutive tasks (Pardo 2006). This will be henceforth referred to as IMITATION. (Though other terms such as ‘convergence’ and ‘accommodation’ are used in the literature to refer to the phenomenon.)

Studies further suggest that social factors can mediate the degree of such imitation. Yu et al. (2013) also observe phonetic imitation of VOT, as well as finding it to be mediated by attitude toward the speaker. Participants perform two blocks of reading a word list out loud, including words starting with voiceless aspirated stops [p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>]. Between the two blocks, participants hear a spoken 1st-person narrative recording. In the recording, the narrator’s VOT of the voiceless stops has been acoustically lengthened. In line with Fowler et al.’s (2003) observations, participants’ own VOTs are longer on average during the second sentence reading block, suggesting that the lengthened VOT of the speaker they listened to has influenced their own phonetic production afterward. However, Yu et al. further observe that this imitation effect is weaker for participants who are found to consider the narrator

less likable in a post-test questionnaire: Their VOTs do not increase as much on average. These observations suggest that attitude toward the speaker can mediate how much the speaker's phonetics affect the listener's own subsequent production. Participants imitate the speaker less when they have a more negative attitude toward them, and more when they have a more positive attitude toward them. If we extend this to loanword adaptation, we should predict a similar outcome: A listener with a more positive attitude toward a speaker first disseminating a loanword will more closely replicate the form they first heard when further disseminating it.

Babel (2010) observes a similar effect, but in this case it is attitude toward the group (rather than the individual) that is found to influence phonetic imitation. New Zealand participants are examined regarding their phonetic imitation of an Australian English speaker. First, participants read a word list so that a baseline plot of their vowel categories can be generated. Then, they perform an auditory repetition task like that used by Goldinger (1998) and Namy et al. (2002): repeating words after hearing the Australian speaker's recorded utterances of them. Before the repetition task, though, participants read a short text describing the Australian speaker. Half of participants are given one version suggesting that the Australian speaker holds positive attitudes toward New Zealand; the other half receive a version leading them to believe the speaker holds negative attitudes toward New Zealand. This is meant to manipulate participants' attitude toward the individual. Finally, Babel's participants also perform an IAT to examine their attitudes regarding Australia(ns) more generally.

Overall, Babel finds that phonetic imitation has occurred: Participants' vowel qualities exhibit a shift in the direction of the Australian speaker's vowels. Furthermore, participants' implicit biases are identified as significant mediators of that phonetic imitation, with stronger anti-Australia biases resulting in weaker imitation of the Australian speaker. However, the speaker description meant to manipulate participants' attitude toward the individual is not observed to significantly affect phonetic imitation. This suggests that bias regarding a group

or region may influence phonetic imitation, besides just attitude regarding the individual. Extending this to loanword adaptation would almost directly mirror Weinreich's (1968, 27) hypothesis: One with a more positive attitude toward the language or group will be more strongly inclined to preserve the form of a loanword disseminated to them from that language or group.

Weatherholtz et al. (2014) also examine an imitation effect, in this case syntactic instead of phonetic. Of relevance here is that they observe political leaning and perceptions of foreignness to mediate imitation. Participants listen to a speech sample manipulated along multiple factors, including a variable syntactic construction: 1) the speaker uses either the dative object construction or the prepositional object construction (the variable tested for imitation) for multiple sentences within a speech sample, 2) the speaker discusses a political topic taking either a liberal or conservative stance regarding it, and 3) the speaker is either a White American representing Mainstream US English, an African American using features of African American English, or a Mandarin-accented non-native speaker of English. These varieties chosen were meant to serve as a continuum of evaluated language standardness. After listening, participants are then instructed to describe certain pictures depicting ditransitive events and therefore eliciting the variable syntactic construction of interest. Results suggest that imitation occurred: In the picture description task, participants are observed to use one construction more if they had heard it more within the speech sample. This imitation is mediated by the other factors manipulated. In a follow-up questionnaire, participants are asked for their own political alignment, and participants are observed to show more syntactic imitation of the speaker when their own political alignment lined up more with the stances they heard in the speech sample. The degree of language standardness is also observed to influence imitation: Participants imitate the White American speaker the most and the Mandarin-accented speaker the least.

Weatherholtz et al.'s observed effects directly relate to the study at hand, seeing that both political alignment and judgments of standardness can influence imitation. It is of



interest here that the judgment of standardness entails a judgment of foreignness, with the most foreign being treated as the least standard. This is relevant to how one interacts with loanwords. Parallel to these results, speakers may not imitate the form they hear as strongly when they consider it foreign (i.e., a loanword). However, syntactic imitation could be different in this respect. In Weatherholtz et al.'s case, participants considered the Mandarin-accented speaker to lack fluency in English. This seemingly led them to resist the speaker's potential influence on their own English syntax. However, for a new loanword uttered by a native speaker of the source language, the foreign-indexed speaker would be the *more* fluent in the source language. Listeners might then give *more* weight to the form they hear uttered by the speaker, therefore more strongly influencing the form they go on to disseminate. Or, at least, this may be the case for some listeners, with the weight being given to the speaker's form potentially mediated by the social factors of interest in this study.

To summarize, there is a large body of evidence suggesting that what one hears can influence their own production. This imitation effect can be mediated by social factors. Those identified as relevant to this study are attitude toward the speaker (Yu et al. 2013), attitude toward the social group (Babel 2010), and political alignment and the association of foreignness (Weatherholtz et al. 2014). It is also worth considering whether such factors are implicit biases or explicit attitudes and stances, both of which have been shown to influence imitation. And, as theorized and observed in Chapter 2, these kinds of social factors appear to be at play regarding the sociolinguistic variation of loanwords. The next section will link the idea of phonetic imitation with loanword adaptation and variation. This is core to the study at hand: testing simulated loanword adaptation as phonetic imitation and examining how this is mediated by social factors which seem relevant to sociolinguistic loanword variation.

### 4.1.2 Loanword variation as mediated source form preservation

These observations of socially mediated imitation are presumed to shed light on how language varieties merge and/or diverge. We assume that people speaking more like each other in these subtle ways on more short-term bases leads to the more distinct language varieties accompanying distinct linguistic communities. Labov (1963), as a seminal example, observes that English speakers in Martha's Vineyard are more likely to use a sound variant associated with that place (the raising of the first part of the diphthongs [aɪ] and [aʊ]) when they hold a more positive attitude toward and stronger alignment with Martha's Vineyard. The subsequent plethora of studies and theories regarding imitation (or 'accommodation', 'convergence', 'audience/referee design', 'adoption') consider this to be a root to how the language varieties of broader communities or generations are formed and delimited (e.g., Trudgill 1972; Bourhis and Giles 1977; Le Page and Tabouret-Keller 1985; Giles et al. 1991; Eckert 2004), though with some debate such as how automatic this is and therefore how solely attributable this may be to the degree of contact between speakers and groups, or to what degree this is attributed to or mediated by attitudes and identities beyond any automatic interlocutor-induced effects. But, broadly, when multiple groups are in contact and we see a certain individual or group speaking more like one other group than another third group, we tend to think that this may reflect something about the relative social dynamics between the speaker(s) of interest and these other groups.

This study assumes the same for loanword variation: that sociolinguistic loanword variation comes about by socially mediated imitation. Loanword variants may come about as a result of social factors mediating the imitation of the source form, leading to a more generalized pattern of the variation between established pronunciations, where those using the more source-like pronunciations may inferrably be those who have a closer relation to the source by shared contact or, degree of contact aside, a stronger alignment and/or more positive attitude toward the source. This all rests, however, on the assumption that established, phonologized loanword variants come about by imitation that can indeed be influenced by

low-level details as well as socially mediated. Phonetic/phonological studies of loanword adaptation and sociolinguistic studies of loanword variation suggest that this is the case.

As mentioned above (Chapter 1, §1.2.3, Kang (2003) and Davidson (2007) observe loanword adaptation to be influenced by the imitation of fine-grained phonetic details, word-final stop releases and excrescent schwa vocoids between consonant clusters respectively. Peperkamp and colleagues (2008) observe a similar phenomenon in which Japanese often inserts vowels after word-final nasals in French-sourced loanwords (e.g., French *Cannes* /kan/ → /kannu/), in spite of Japanese native words readily allowing nasals in coda position (e.g., ‘rice’ /gohan/). They conclude this is due to the fact that French audibly releases word-final nasals phonetically (*Cannes* [kan̩]) while Japanese does not (‘rice’ [gohan̩]), further supported by the observation that English-to-Japanese loanwords do not tend to exhibit this pattern (e.g., *salaryman* [sæləri mæn̩] → /sarari:man/). Such observations suggest that phonetic detail can influence loanword adaptation. Speakers may be particularly attentive to and influenced by lower-level phonetic details when adapting and disseminating a loanword. And seemingly, they are strongly inclined to preserve such details.

There are also sociolinguistic studies examining loanword variation. They each discuss this variation as the degree of (posited) source form preservation, with certain groups or styles preserving the source form more than others. This could be the result of such groups’ replication of loanword source forms being mediated by social factors during adaptation. As discussed above (Chapter 1, §1.3), Boberg (1997, 1999) observes the variation of ‘foreign (a)’ between more source-like (or at least positedly more source-like) [ɑ] and less source-like [æ] to pattern with linguistic security and as an apparent change in progress. And, Thomason (2001) describes how Islam-related Arabic loanwords undergo unnativization in formal styles of Turkish, preserving Arabic sounds not considered available in normal Turkish phonology. Van Oostendorp (1997) observes a similar phenomenon in Dutch. In more formal styles of Dutch, loanwords’ vowel qualities can be preserved in unstressed syllables, in spite of Dutch’s inclination to neutralize unstressed vowels: e.g., *democratie* as [dèmokrətí] in a

more casual style but as [dèmokratí] in a more formal style (219). Van Oostendorp formalizes this in an Optimality-theoretic framework as an upward shifting of the relevant faithfulness constraints to achieve the formal state's grammar. This therefore directly formalizes the observed variation as stylistically variable preservation of the source form, where more formal styles induce a stronger faithfulness to the source form.

Poplack, Sankoff, and Miller (1988) also show that loanword variation is analyzable as preservation of the source form. Additionally, they demonstrate the temporally gradual nature of loanword adaptation, which further supports considering the variation of established loanwords a result of social effects on the path of adaptation. They examine English loanwords in Canadian French. When they examine loanword *usage*, they find that communities with stronger English contact use loanwords more frequently and use newer, more innovative ones. And between communities of similar degrees of contact intensity, variation is still observed. They interpret that these communities have differing attitudes toward Anglicisms and regarding the purity of French. These interpretations relate to the factors of interest in this study of source-directed attitude and/or broader language contact ideology. Poplack and colleagues also look at speakers' proficiency in English, finding higher English proficiency to also correlate with loanword usage; however, community is found to serve as a stronger predictor. When they examine "nativization" (i.e., the replacement of French sounds with English sounds), they find that Canadian French speakers with higher levels of English proficiency nativize English loanwords to a lesser degree. They also find that older and more frequent loanwords tend to be more nativized.

These observations by Poplack and colleagues suggest that what may be synchronically seen as loanword variants across social dimensions or groups may more accurately be considered different degrees or stages of adaptation from the source form. In an apparent time approach, this suggests that loanword adaptation is a gradual path, and the communities and speakers using more source-like variants seem to be less far along that path, more strongly preserving the source form. However, as observed in Chapter 2, some loanwords ap-

pear to be quite established yet still categorically variable between more and less source-like variants. This calls for a revised formulation of how socially mediated loanword adaptation can lead to loanword variation: Social factors may mediate how strongly one speaker or group preserves the source form and resists full adaptation to the borrowing language's sound system. However, weaker imitation may also lead adaptation toward a different path, toward a form that is fully complicit in the borrowing language's sound system but even less resemblant of the source form than a full adaptation that would have been reached by maximal imitation of the source form.

In a more experimental approach, Lev-Ari and Peperkamp (2014) provide online evidence for socially mediated loanword adaptation. French speakers are exposed to [ʤen:a], a nonce word framed as being Italian. In some trials, *genna* is illustrated as a gelato, a domain in which Italian products carry high prestige. In others, it is illustrated as a beer, a domain in which Italian products carry low prestige. Participants are more inclined to preserve the Italian source form [ʤ] (rather than adapting it to French [ʒ]) when the word is framed as representing a gelato and therefore pertains to a semantic field in which they associate Italian with more prestige. Therefore, source form preservation is stronger when speakers hold a more positive attitude toward the source language, while that attitude is mediated by semantic associations.

The study at hand intends to examine loanword adaptation similarly, as the socially mediated replication of (posited) source forms of experimental nonce words framed as loanwords. In contrast to Lev-Ari and Peperkamp, this study employs a phonetic imitation paradigm like the studies mentioned above (Babel 2010; Yu et al. 2013), examining how much certain social factors may enhance or mitigate the imitation of phonetic detail, instead of posited segmental representations. Furthermore, this study teases loanword framing apart from phonological non-nativeness: Rather than testing for the imitation of a sound that is not considered a member of borrowing language's native inventory (like [ʤ] in an Italian-to-French borrowing), it is ensured that the exposure form tested for imitation (i.e., the posited

source form) is natively attested and licit in the sound system of the borrowing language.

## **4.2 Methods**

The utterances analyzed in this study come from the first major set of tasks in the larger study which included the tasks described in Chapter 2: the elicitation of established variable loanword utterances and a follow-up Implicit Association Test (IAT) and Likert agreement questionnaire to measure the social factors of interest for each participant. This part of the study is designed to auditorily expose participants to nonce words within three short stories, elicit subsequent utterances of those nonce words in short sequels they read aloud, and examine how much phonetic imitation occurs (i.e., how much the auditory exposure influences participants' subsequent pronunciations). Nonce words are framed differently across the three stories so that participants treat them as either an unfamiliar English word, an Iraq-sourced loanword, or an Indonesia-sourced loanword. Then, the same social factors analyzed as predictors of established loanword variation in Chapter 2 are also analyzed as predictors of nonce word imitation, hypothesizing that those more likely to use more source-like variants of established loanwords are also those who more strongly imitate the forms of nonce loans encountered for the first time.

### **4.2.1 Stimuli**

#### **4.2.1.1 Phonetic variables**

Two PHONETIC VARIABLES were chosen to manipulate and examine. Each was manipulated to represent two oppositional pronunciations, or VARIANTS. Participants' productions could then be acoustically examined to test which variant their pronunciation was more similar to. Attention was paid to using variables in which either potential variant is already attested in the phonetics and phonology of American English. This was meant to purposefully tease apart phonetic imitation from what could be considered phonological adjustment to accommodate an otherwise non-native sound form, focusing on the first. If participants were

exposed to a form that might be considered phonologically illicit in American English, it would be harder to know whether attenuated influence of the exposure form was due to less imitation or more phonological constraint and/or articulatory difficulty in replication.

One variable manipulated was VOWEL REDUCTION. For this variable, the first vowel in disyllabic nonce words with final stress was manipulated between [ɛ] and the more reduced [ə]: e.g., [dɛníɪ]~[dəníɪ]. This phonetic variation of vowel reduction in word-initial syllables is encountered within American English. For example, similar variable vowel reduction is encountered in words like *lethargic* [lɛθáɪɟɪk]~[ləθáɪɟɪk]. It is also found in established loanwords in American English such as *petit(e)* [pɛtít]~[pətít].

The other variable manipulated was GLIDING. For this variable, the final sequences of nonce words ending orthographically in <ia> were manipulated between [iə] hiatus and glide-vowel [jə] forms: e.g., [hɛníə]~[hɛnjə]. This is also a variable already encountered within American English. For example, there is similar variation found in the word *gardenia* [gɑɪdíniə]~[gɑɪdínjə], as well as placenames such as *Armenia* [ɑɪmíniə]~[ɑɪmínjə] and *Estonia* [ɛstóniə]~[ɛstónjə].

For each variable, three nonce words were constructed, one for each story participants would listen to. The six target nonce words can be found in Table 20, along with their orthographic representations across framings (which will be discussed in the next section) and filler nonce words. For each stimulus, a participant would hear one variant or the other to then test how much that exposure influenced their subsequent production when reading a sequel aloud.

#### 4.2.1.2 Framing

Participants were exposed to the nonce stimuli within three short stories they listened to. Exposure included an audio-recorded telling of the story accompanied by illustrations. Multiple aspects of these exposures were manipulated to create the three different framings: US (non-loan), Iraq (loan), and Indonesia (loan). The two non-US loan framings were chosen in

part to keep constant the fact that each place is a Muslim-majority country, while simultaneously teasing apart political charge of the place: Iraq is much more frequently discussed in US politics in a charged fashion; furthermore, it is expected that participants do not as strongly associate Indonesia with Islam, in spite of the fact that it is also a Muslim-majority country. These two separate foreign framings are therefore intended to allow for the examination of source-directed attitude as an influence of replication, expecting that participants may have more charged attitudes regarding Iraq, while considering Indonesia as foreign but without as potentially negative an attitudinal charge. It is therefore predicted that participants with stronger anti-Iraq attitudes (explicitly expressed or measured in the form of implicit associations) will less strongly replicate the sound forms of nonce loans they consider Iraq-sourced compared to their imitation of those they consider Indonesia-sourced.

A different SPEAKER was recruited to record the stories for each framing, all adult males. A white native speaker of American English was used to record the US framing. An Iraq native who speaks Arabic as a first language and learned American English as a second language was used to record the Iraq framing. An Indonesia native who speaks Javanese as a first language and learned American English as a second language was used to record the Indonesia framing. This was done so that each speaker could be sociophonetically indexed by the listener with the framed region. Recordings were focus-grouped amongst some phoneticians to ensure that F0 and voice quality did not seem drastically different to listeners, as well as to ensure that speakers' degree of accentedness and difficulty in comprehension was not notably different between the Iraq and Indonesia framings.

The story SETTING was altered per framing. The place where the story took place was explicitly mentioned at the beginning of each story, and the illustrations for each started with a map of the country mentioned to further highlight the setting. The narrative was set in 1st person, with the speaker identifying the place as his home region to further connect him and the incorporated nonce words to that region.

For the loan framings, there was always a MENTION OF THE "LOCAL LANGUAGE"



(though never specific) in the early part of the story, before the target nonce words were first introduced. For example, in one story an American colleague is joining the narrator to assist on a mission. The line introducing the character reads: “Allison was my assistant who traveled with me and helped me prepare[, and I translated between her and the local residents].” The final bracketed part was only included in the Iraq and Indonesia framings.

Finally, illustrations were labeled 1) so that nonce words could be easily associated with the entities they represented, and 2) so that their orthography was familiar by the time participants encountered them during the sequel reading task. For these illustration labels, the ORTHOGRAPHY was manipulated across framings so that participants would form different phonographic associations (i.e. spelling-to-sound mappings) to further prime a sense of nativeness-vs.-foreignness. Table 20 shows the nonce words, including the variants heard and the different orthographic representations for each framing.

VARIABLE	NONCE WORD	FRAMING		
		US (NATIVE)	IRAQ (LOAN)	INDONESIA (LOAN)
REDUCTION	[dɛníɪ] ~ [dɛníɪ]	deneer	denihr	denír
	[sɛmév] ~ [sɛmév]	semave	semev	semév
	[zɛnúl] ~ [zɛnúl]	zenool	zenuhl	zenúl
GLIDING	[hɛníə] ~ [hɛnjə]	hennia	hehnia	hénia
	[númiə] ~ [númjə]	noomia	numia	númia
	[ʃíniə] ~ [ʃínjə]	sheenia	shinia	shínia
FILLERS	[klot]	clote	kloht	klót
	[kjum]	cume	kjum	kyum
	[gle]	glay	gleh	glé
	[gʒdʒi]	gurgee	gurdzhi	gurji
	[sláksi]	sloxy	slaqsi	slaksi
	[sándo]	sondow	sandoh	sándo

Table 20: Nonce stimuli

Below (Figure 17) is the same illustration panel as presented across the three different framings. This includes the orthographic labeling of two nonce words, both target and filler, representing a kind of flower and a kind of flute respectively. Appendix F provides the full

scripts for all stories and sequels across all framings.

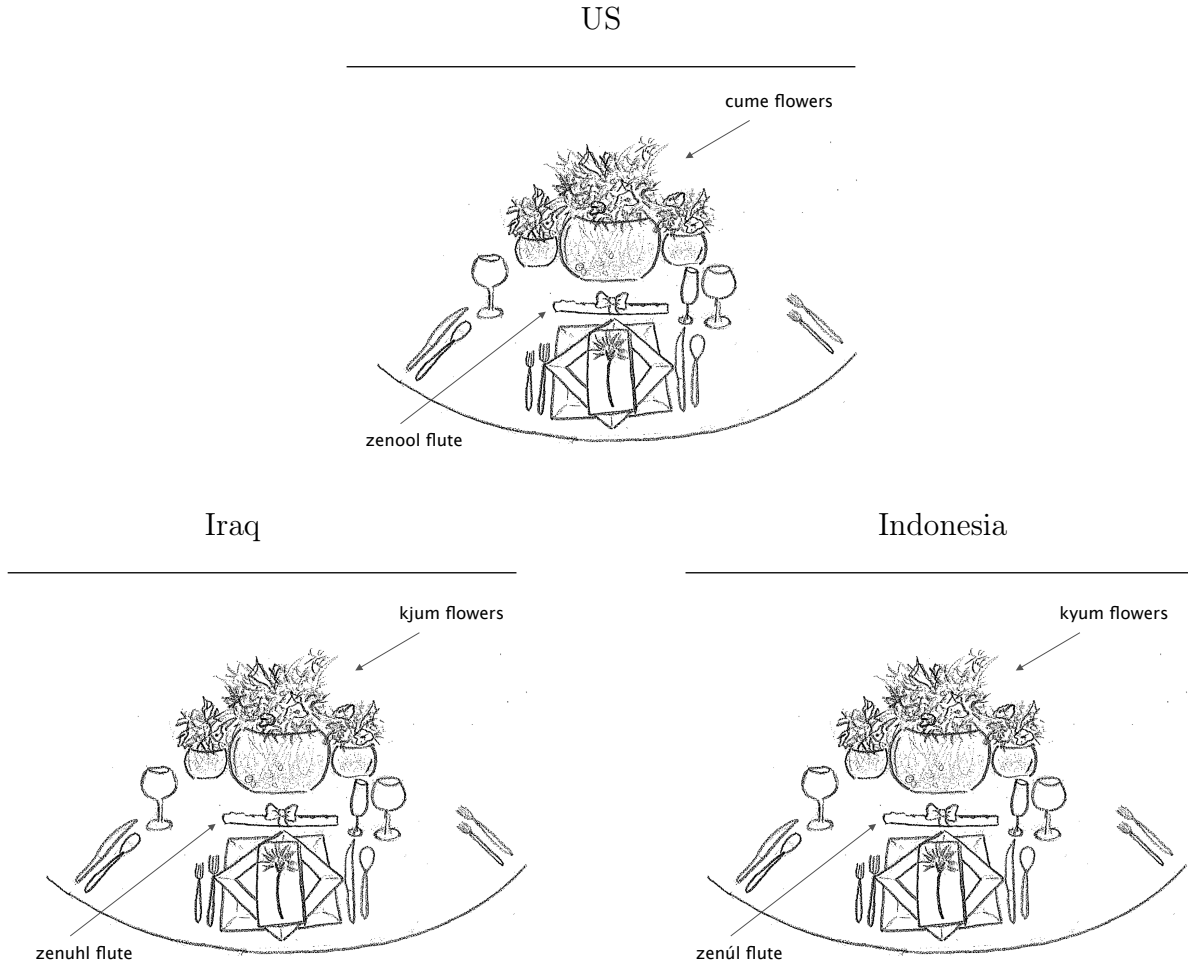


Figure 17: Sample illustrations including orthographic labeling

#### 4.2.1.3 Preparation

Three different stories were written with corresponding sequels. The central THEME of each pair will be used to refer to them: Exploring, Visiting, and Wedding. Illustrations were then drawn by a skilled colleague for each theme's story. (Other than label orthography and the initial map setting, illustrations remained constant across framings within each theme.) Each story + sequel pair incorporated one target nonce stimulus per variable and two filler nonce stimuli. These assignments between theme and nonce stimuli remained constant. Within each story and each sequel, each nonce stimulus appeared twice. Attention was paid

to evenly spacing the target nonce stimuli from each other and alternating them. Within each sequel, attention was also paid to controlling the environment in which nonce words occurred: placing them in a prosodically prominent position near the beginning of a sentence, and controlling the following segment for the word-final GLIDING variable (all were followed by a word beginning with a voiceless labial obstruent). For each theme, the US, Iraq, and Indonesia framings were then created by modifying the factors of setting, mention of the “local language”, and orthography as described above. This resulted in 9 story + sequel pairs: 3 themes  $\times$  3 framings.

Recording for the stories was done at the NYU Phonetics and Experimental Phonology laboratory. The speaker was seated in a sound-attenuated booth with me, the researcher, and recorded using a Shure SM35 headworn microphone connected to a Zoom H4n digital audio recorder (44.1kHz, 24bit sampling). The speaker was trained regarding the phonetic variables of interest. Each speaker read the stories from all three themes with the text in their respective framing; however, all three speakers saw the US-framed orthography of the nonce words. While the speaker read the story, he would be asked to repeat each sentence containing a target nonce word multiple times using one pronunciation variant, then multiple times using the other. Attention was paid to ensuring that there were takes for each variant, audibly distinct from each other and with a similar intonational contour, as well as ensuring that any disfluencies were removable.

A final version for each story was created using the acoustic analysis and editing software Praat (Boersma and Weenink 2015). First, the best take of each utterance was captured, removing any disfluencies and aiming for a similar rate of speech and narrative tone across them all. Then the vocalic material of interest in each target nonce word was substituted with that from the optimal utterance of the other variant (taking into account both accurate phonetic realization of the variant and prosodic fit with the surrounding context). Any further manipulation was performed to ensure that the variants sounded

natural and distinct.<sup>5</sup>

The manipulated variables were then cross-balanced, resulting in the four possible combinations of each variant: [ə]+[iə], [ə]+[jə], [ɛ]+[iə], [ɛ]+[jə]. This resulted in a total of 36 different versions of the exposure stories: 3 themes × 3 framings × 4 variant combinations. All audio files were intensity-scaled using Praat. Then, using iMovie 10.1.2 running on Mac OSX 10.11, the 36 different story recordings were paired with their respective labeled illustrations and saved as .m4v video files.

#### 4.2.2 Procedure

As described in Chapter 2, participants were seated in a sound-attenuated area. On the desk in front of them was a computer screen, keyboard, and mouse. Participants' speech was recorded using a Shure SM35 headworn microphone connected to a Zoom H4n digital audio recorder (44.1kHz, 24bit sampling). The entire experiment was developed and administered via the PsychoPy software (Peirce 2016).

Before the main task exposing participants to the nonce words of interest in short stories, participants first performed a sentence reading task. A sentence was displayed, which the participant was prompted to read aloud; the participant would then press a button to advance to the next sentence. Key words within these sentences were used to elicit participants' utterances of the categories of interest to the phonetic imitation analysis: for the vowel reduction variable, [ɛ] and [ə]; for the gliding variable, [jə] and [iə]. These utterances were acoustically analyzed to serve as participant-specific baseline measurements and test which category a participant's nonce word utterance more closely resembled (and, therefore, whether their utterance more closely resembled the variant they were exposed to auditorily in the preceding short story). Acoustic measurement and analysis is more fully detailed below: §4.3.2.1). Table 21 lists these 14 key words along with some example sentences. Attention was paid to placing words in early, prosodically prominent positions within each

---

<sup>5</sup>For example, in the GLIDING condition the preceding vowel and nasal segment's duration may have been altered to achieve the disyllabic percept of the [jə] variant and trisyllabic percept of the [iə] variant.

sentence. For the gliding variable, the following word in the sentence always began with a voiceless labial obstruent to match the following environment in sequels eliciting participants’ utterances of nonce words with this variable. In terms of presentation order, variables (and variants of each variable) were alternated between and evenly spaced. Key word stimuli were then further spaced by the addition of filler sentences. The order of presentation was kept consistent across all participants. The task elicited the utterances of 26 sentences in total.

VARIABLE	VARIANT	WORDS
REDUCTION	[ɛ]	centipede, demonstrate, dentist, seminar
	[ə]	demolish, deny, denounce, development
GLIDING	[jə]	California, Kenya, pneumonia
	[iə]	Bosnia, hernia, millennia
EXAMPLE SENTENCES		The dentist was late to the appointment. She’s going to Kenya for a conference next week.

Table 21: Key words elicited for baseline measurements

The following task was the phonetic imitation task exposing participants to nonce words and eliciting their subsequent reproduction of them. Participants were instructed that they would listen to three short stories presented with illustrations and that after each story they would read aloud a sequel presented to them on the screen. They would press a key on the keyboard to advance and begin the first story. After it finished, the text of the sequel then automatically appeared for participants to read aloud. After reading the sequel, participants would then press a key on the keyboard to advance to the next story + sequel trial.

Every participant performed three such trials. Table 22 shows two examples of what a participant may have experienced. All three themes and all three framings were represented across the three trials. Across all participants, the ordering of theme was kept constant: 1) Exploring → 2) Wedding → 3) Visiting. All permutations of framing order were assigned randomly, while evenly, across participants. Nonce stimuli remained constant within each theme: e.g., the Exploring theme always used the nonce stimuli [sɛmév]~[səmév] and

[hénjə]~[hénia]. Orthography of nonce stimuli was dictated by framing, in line with Table 20: hence the nonce stimuli for participant X in trial 1 are orthographically presented as <semave> and <hennia>, because they are in the US-framed trial. Every participant heard three of the four possible variant combinations across the three trials. This ensured that every participant encountered both variants of each variable. Like framing ordering, this assignment was also randomized while balanced across participants.

PARTICIPANT X	1		2		3	
TASK	story	sequel	story	sequel	story	sequel
FRAMING	US		Indonesia		Iraq	
VARIANTS HEARD	[ə] + [jə]		[ə] + [iə]		[ɛ] + [jə]	
ORTHOGRAPHIES	semave, hennia		zenúl, shínia		denihr, numia	
THEME	Exploring		Wedding		Visiting	

PARTICIPANT Y	1		2		3	
TASK	story	sequel	story	sequel	story	sequel
FRAMING	Indonesia		Iraq		US	
VARIANTS HEARD	[ɛ] + [iə]		[ə] + [iə]		[ɛ] + [jə]	
ORTHOGRAPHIES	semév, hénia		zenuhl, shinia		deneer, noomia	
THEME	Exploring		Wedding		Visiting	

Table 22: Example procedure for two (hypothetical) participants

The tasks following this are described in Chapter 2. After finishing the sequel + story trials, participants performed a sentence reading task focused on eliciting their utterances of established variable loanwords, the variation of which were the focus of Chapter 2. Participants then performed an IAT to test for possible negatively biased associations toward Iraq relative to those held toward Indonesia: i.e., testing whether and to what degree participants have faster reaction times sorting stimuli when associating Iraq-indexed images with negative words via a shared response key vs. associating Indonesia-indexed images with negative words. Finally, participants completed a 7-point Likert agreement questionnaire to generate an index of where each participant identified along social factors of interest: political identity, Iraq directed attitude (explicit; still relativized to attitudes toward Indonesia), language contact ideology, globalist/nationalist ideology, mobility, and prescriptivism. (See §§2.3.2.1,

2.3.3.1 of Chapter 2 for a more detailed reminder of the contents and procedure of these two tasks and the generation of their resultant measurements.)

### 4.3 Results

Like in Chapter 2, each participant's scores from the IAT and Likert questionnaire are examined as predictors of the linguistic behavior elicited in this task. It is predicted that those more likely to use more source-like variants of established loanwords are those who more strongly imitate the exposure form of nonce loanwords. In terms of the phonetic variables manipulated, imitation here means that when someone heard the [ɛ] exposure variant of the vowel reduction variable, their utterance should more strongly resemble their own [ɛ] category than their [ə] category as compared to if they heard the [ə] exposure variant; or, when someone heard the [jə] exposure variant of the gliding variable, imitation would mean their utterance should more strongly resemble a [jə] realization than a [iə] realization as compared to if they heard the [iə] exposure variant. Weaker imitation would mean that the exposure variant a participant heard had less influence on their pronunciation of the nonce word: i.e., they might still use a [ə] pronunciation in spite of having heard the [ɛ] variant during exposure, or a [iə] pronunciation in spite of having heard the [jə] variant during exposure. (And, as will be observed below, [ə] and [iə] are the apparent default preferences for each variable.)

The two different loan framings allow for the examination of source-directed attitudes as a mediator of this imitation, predicting that those with stronger anti-Iraq biases (implicit and/or explicit) will show stronger imitation of nonce words framed as Indonesia-sourced loans than those framed as Iraq-sourced loans. The US framing allows for the examination of loan status as an influence on imitation: It's possible that those who are, say, more globalist or politically more liberal-identifying will show stronger imitation of loanwords, no matter their source, relative to their imitation of unfamiliar words that they do not consider (or are at least less primed to consider) to be of foreign origin.

### 4.3.1 Categorical imitation results

Before analyzing the imitation of the phonetic variables manipulated during exposure (gliding and vowel reduction), a broader analysis outside of this envelope is motivated. When segmenting participants' utterances of nonce words, considerable variation was observed in the imitation of sounds not phonetically manipulated at all. For example, the filler nonce word presented consistently within and across framings as [sláksi] was observed to be reproduced as [sláski], [slákski], [sléksi], etc. This also extended to the target words of interest for phonetic imitation. For example, the nonce word manipulated between [ʃínjə] and [ʃíniə] variants was reproduced by some participants with clear divergences from either exposure form: e.g., [ʃínjə], [ʃəniə], [ʃənáriə]. It was considered that the same hypotheses regarding socially mediated imitation might apply to this variation as well: that participants whose reproductions involved these divergences from the exposure form might be those socially predicted to do so, such as those who identify as more nationalist or politically conservative, or those who hold a more negative attitude toward the framed source.

This variation was analyzed with an approach similar to that of the variation of established loanwords undertaken in Chapter 2. For each utterance, the number of categorical divergences was tallied. This included sound substitutions (e.g., [klot] → [klat]), sound insertions (e.g., [zənúl] → [zənjúl]), sound deletions (e.g., [gle] → [ge]), metatheses (e.g., [sláksi] → [sláski]), and stress misplacement (e.g., [sándo] → [səndó]). For example, an utterance of the nonce word of exposure form [sláksi] as [sléski] would be tallied as 2 divergences: the changing of [ɑ] to [æ] and the metathesis of [ks] to [sk]. A percept of either variant for the vowel reduction and gliding variables was not counted, as this is analyzed phonetically next (§4.3.2). Then, for each participant, an average number of divergences was calculated per word.

Before proceeding to statistical analysis of the effects of the social factors, some further attention should be paid to how aspects of the condition assignment may have played a role. Figure 18 shows the average number of divergences per word across all participants, grouped



by trial number (horizontally) and separated by framing (vertically). As is apparent, some words show different relative averages: e.g., [kjum] from trial 2 shows a lower average of divergences across all framings relative to the other nonce words in that trial. However, framing clearly also seems to have an influence. Broadly speaking, there is a lower average number of divergences (though not that of 0) in the US framing. This effect of framing also appears to interact with word. For example, the reproduction of [gʒɔ̃ʒi] shows a much higher average number of divergences in the Iraq framing than in the Indonesia framing. This, however, should be interpreted with caution. Recall that orthography was manipulated along with framing (Table 20). This increase in divergences from the [gʒɔ̃ʒi] exposure form in the Iraq framing is surely due to the fact that its orthographic representation in this framing was <gurdzhi>, which led some participants to produce it as [gʒɔ̃ʒi] (cf. <gurjee> in the US framing, <gurji> in the Indonesia framing). Much of this divergence may therefore be attributable to phonography (i.e., spelling-to-sound mapping). The attribution of such effects to phonography does not necessarily mean that this variation cannot be socially mediated and, therefore, socially meaningful. Since all participants did hear multiple auditory exposures of the source form during the story presentation, alongside simultaneous display in the illustrations of the orthographic representation, some were still clearly able to maintain a faithful replication of the source form during the sequel reading task and override any phonographic biases that may have been activated. Given this, we may therefore more accurately consider this imitation as imitation in spite of phonographic bias. However, the phonography of all sounds was not tightly controlled or balanced across framings (as was the case for the [ə]~[ɛ] and [jə]~[iə] variables). These results should therefore be treated with caution, recognizing the potential conflation of framing with particularly difficult phonographic associations.

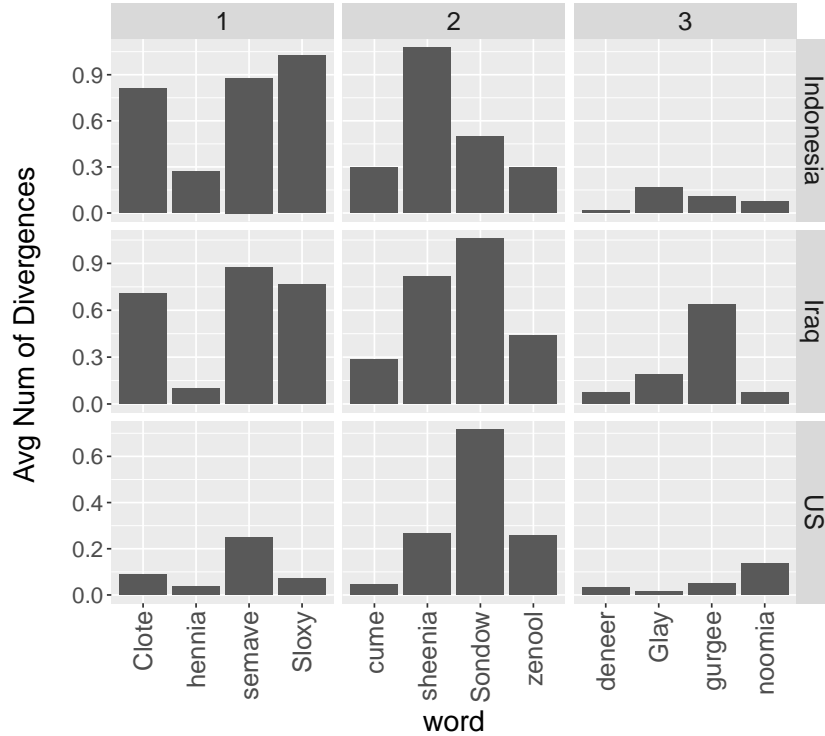


Figure 18: Categorical imitation: word  $\times$  framing

Average number of categorical divergences per nonce word, grouped horizontally by framing and grouped vertically by exposure trial. (Orthography provided per word is that for US framing.)

Statistical analysis of these findings resembles that performed in Chapter 2: the generation of mixed-effects models with predictors of interest and a step-up comparison of those models to test for significance of the effects of those predictors. In the following statistical reporting, effect coefficients ( $\beta$ ), standard deviations ( $\sigma$ ), and  $t$ -values ( $t$ ) come from the final model including all fixed effects identified as significant contributors to model fit;  $p$ -values ( $p$ ) come from a Chi-square ANOVA test between that full model and a model excluding the effect of interest (Winter 2014). The base model for this analysis, before incorporating social factors of interest, considers the effects of word and framing. As discussed above, some words overall show higher averages of divergence from the exposure form, motivating a random intercept per word. As also discussed above, some words show higher averages of divergence particular to the framing in which the participant was exposed to them (likely due to phonography). This motivated a random intercept per word + framing combination. The

effect of framing was included as a base fixed-effect, considering that different framings also appear to have different averages of divergence and that this is relevant to the hypotheses at hand. (This effect is indeed identified as significant, as shown below, mainly as an effect of loanword status: Both Iraq-framed nonce words and Indonesia-framed nonce words show higher averages of divergence than US-framed nonce words, while it is notable that nonce words which are Iraq-framed still show a slightly stronger tendency toward divergence.) Modeling then proceeded to consider each social factor of participants as an additional fixed effect, identify the strongest and, if the strongest significantly improved model fit, subsequently test for any additional predictors that significantly improve model fit even further. With the consideration of each social factor, an interaction with the effect of framing was also considered and tested for significance. The results of the final model converged upon are provided in Table 23 below. Appendix G provides a more detailed summary of the step-up modeling performed.

FACTOR	INTERACTION (if any)	$\beta$	$\sigma$	$t$	$p$
(intercept)		.163	.093	1.754	
framing = Indonesia		.297	.089	3.352	.00091
framing = Iraq		.339	.089	3.834	
globalist/nationalist alignment		.038	.015	2.54	.011
Iraq-directed attitude (Likert)	framing = US	-.009	.024	-0.404	.023
Iraq-directed attitude (Likert)	framing = Indonesia	.071	.024	2.901	
Iraq-directed attitude (Likert)	framing = Iraq	.023	.024	0.942	
political identity	framing = US	.025	.025	0.998	.053
political identity	framing = Indonesia	.039	.025	1.594	
political identity	framing = Iraq	-.049	.025	-1.969	

Table 23: Step-up comparison results: Categorical imitation

Results of final linear mixed-effects model after step-up comparison of independent variables as predictors. Each factor's  $p$ -value comes from a Chi-square ANOVA test between the full model and a model excluding the component of interest. ("Intercept" = US framing.)

In parallel with the variation of established words observed in Chapter 2, the factor of a speaker's globalist/nationalist alignment was identified as the strongest social factor predicting exhibiting an effect on imitation. Figure 19 plots all speakers' average number of divergences per word, with data points grouped by the framing in which the speaker was

exposed to that word. This effect is in the expected direction: Those with a more globalist alignment are those who show stronger imitation of the exposure forms of nonce loanwords. Interestingly, however, this effect appears to hold across the board and the inclusion of an interaction of this effect with that of framing is not motivated. As is visible in Figure 19, this effect is quite parallel for nonce words framed as loans and those framed as non-loans: Those with a more globalist alignment appear to be stronger imitators of nonce word exposure forms in general, no matter their loanword status.

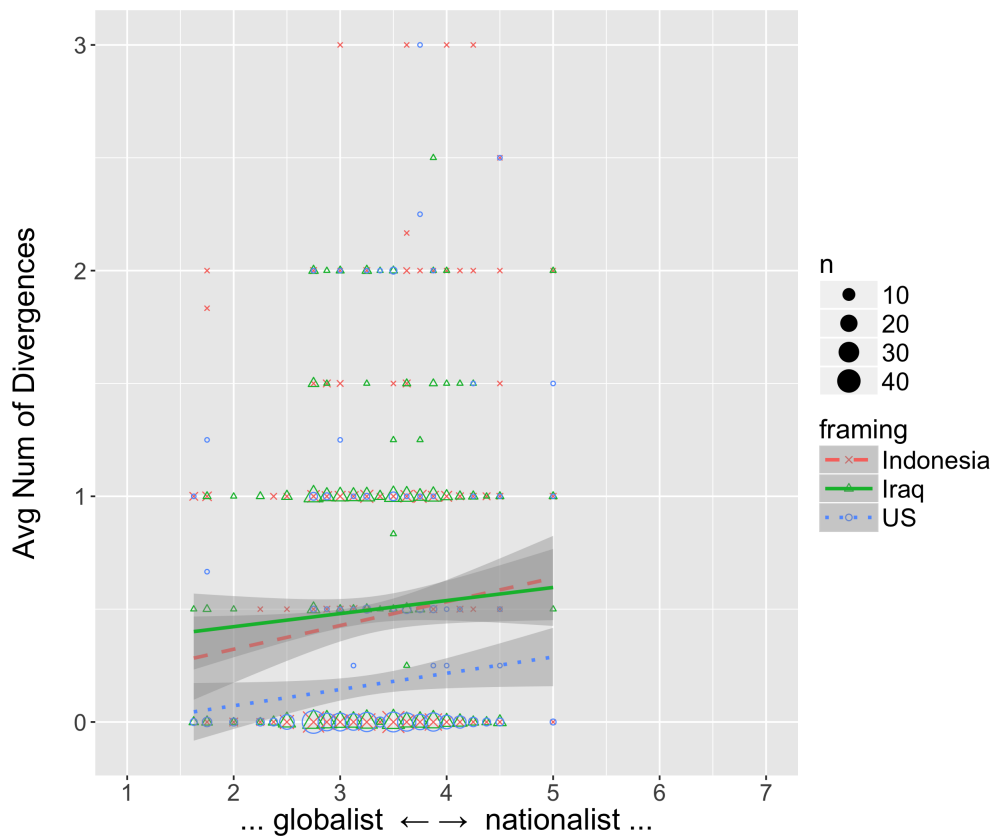


Figure 19: Categorical imitation: globalist/nationalist alignment

The y-axis represents the average number of divergences per word per speaker, the x-axis is the composite globalist/nationalist alignment score of the participant based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Point sizes represent the number of data points that would otherwise be overlapping.

One additional factor is identified as having a significant effect: an interaction of the effect of explicit place-directed attitude with framing. The direction of this effect, however, does not pattern as would be expected. As is apparent in Figure 20, those who showed a

stronger disfavoring of Iraq in their responses to the Likert questionnaire show more divergences from the source forms of nonce loans framed as Indonesia-sourced.

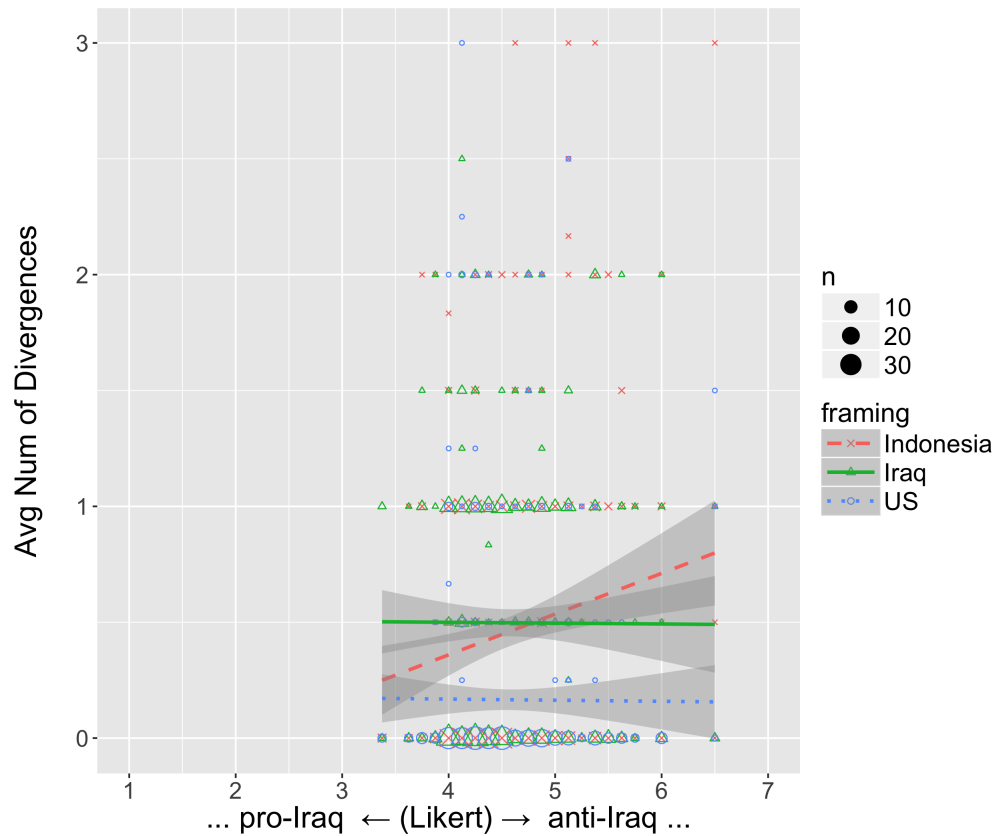


Figure 20: Categorical imitation: place-directed attitude

The y-axis represents the average number of divergences per word per speaker, the x-axis is the composite score of the participant's relative anti-Iraq attitude based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Point sizes represent the number of data points that would otherwise be overlapping.

When further checking the data, this appears to be attributable to the utterances of multiple nonce words by multiple speakers: i.e., the cluster of Indonesia-framed tokens in the upper right of Figure 20 do not all come from one speaker or from one nonce word. This therefore does not seem easily explainable by any particular outliers along either random dimension. A clear interpretation and conclusion regarding this effect may not be reachable at this time. However, as mentioned above, it is possible that this is a result of the conflation of framing with phonographic association. As is notable in Table 20, one way orthography was used to enhance the sense that nonce words were Indonesia-sourced was by the inclusion

of the <´> diacritic; orthography in the Iraq framing, however, did not use any diacritics. It is possible that the presence of diacritics was particularly jarring to participants, therefore making the replication of the source form in the Indonesia framing more difficult. Why this is especially more difficult for those with stronger anti-Iraq attitudes is not entirely clear. It could be that this effect is explained in part by the fact that language contact ideology is the social factor which shows the strongest collinearity with this factor: Those who show stronger anti-Iraq attitudes are also those who express less receptiveness regarding language contact ( $r=.31$ ). And, the above interaction with framing is similarly pronounced when examined visually along the factor of language contact ideology in Figure 21. So, perhaps not being receptive to words or word forms that feel more foreign leads one to have more difficulty processing orthographic forms that are immediately indexed as foreign (and perhaps to an especially high degree) due to the presence of a diacritic. Though, again, this may only be conjecture at this point; we must remember that anti-Iraq attitude was indeed identified as the factor more strongly capturing this variation, while a model replacing the effect of anti-Iraq attitude with that of language contact ideology fares significantly worse at fitting the data ( $p<.0001$ ).

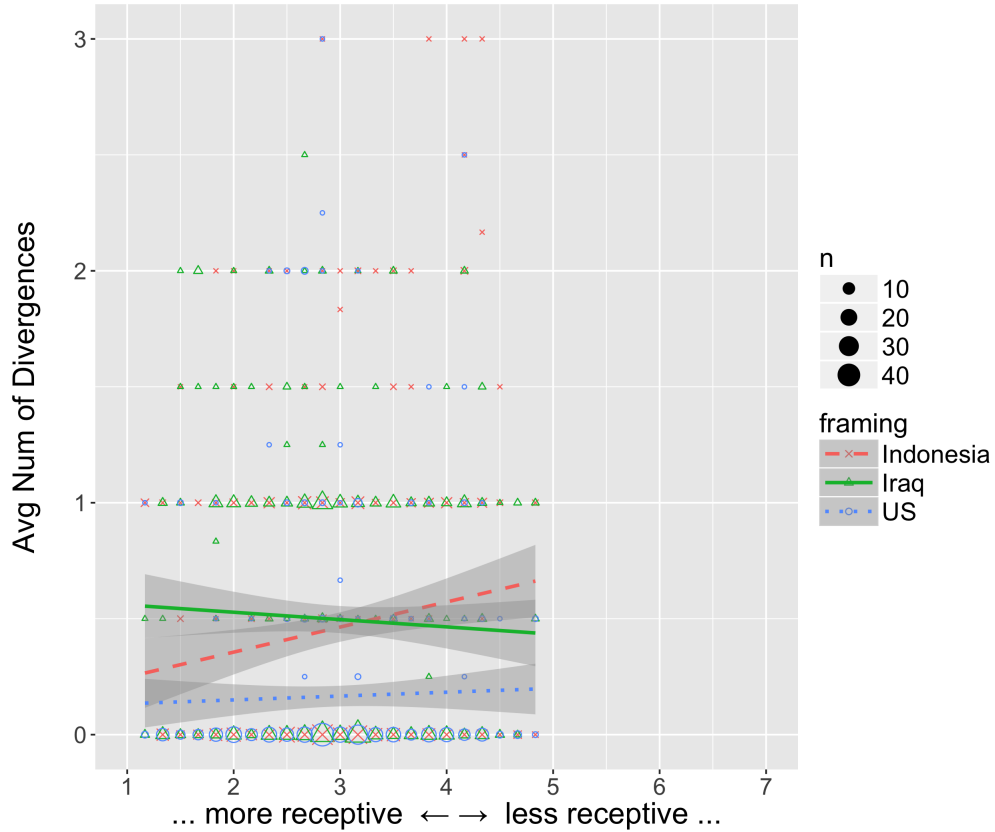


Figure 21: Categorical imitation: language contact ideology

The y-axis represents the average number of divergences per word per speaker, the x-axis is the composite score of the participant’s receptiveness regarding language contact based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Point sizes represent the number of data points that would otherwise be overlapping.

While no other effect is identified as a significant further contributor to model fit, the interaction of political identity with framing is identified as a trend. This effect, however, is also not entirely clear. The main component of this interaction contributing to model fit is that regarding the Iraq framing, where those who are more liberal-identifying show more divergence from the source forms of Iraq-framed nonce loans, as visualized in Figure 22.

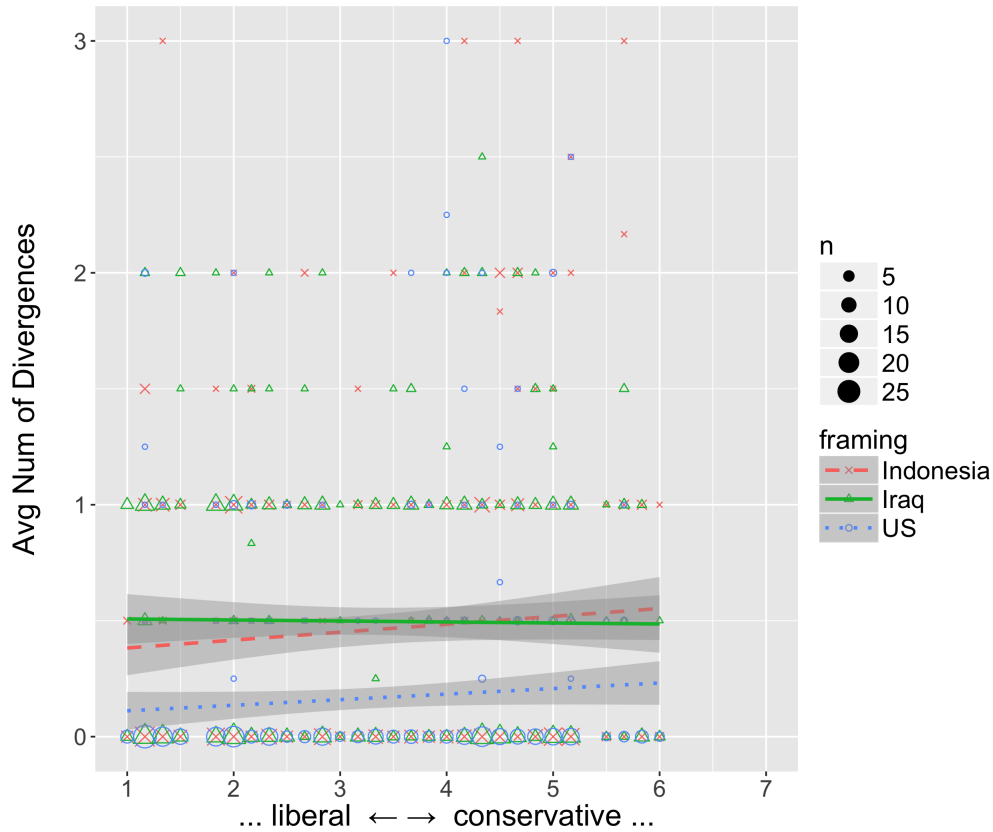


Figure 22: Categorical imitation: political identity

The y-axis represents the average number of divergences per word per speaker, the x-axis is the composite score of the participant’s receptiveness regarding language contact based on their responses to the Likert questionnaire, and the line is a curved generalized linear model with a 95% confidence interval. Point sizes represent the number of data points that would otherwise be overlapping.

This effect is notably in the reverse direction than we would expect based on the discussion in Chapter 2, if interpreting the pattern of political identity to fall out from the correlation of anti-Arab or -Middle East attitudes with conservative political identity (Nosek et al. 2007; Arab American Institute 2014). However, it is also notable that explicit anti-Iraq attitudes in the Likert questionnaire show a relatively weak collinearity with political identity amongst participants represented in this dataset ( $r=.068$ ). Therefore, this dataset is teasing political identity apart from source-directed attitude. But, given that, this observation still patterns against what we would expect of political identity, in and of itself, based on Hall-Lew et al.’s (2010; 2012) findings that those more liberal-identifying use the more source-like pronunciation of Arab-sourced *Iraq*. We may not be able to reach a clear answer



regarding this observation at this time. As conjectured regarding the effect of place-directed attitude, this could also be explainable by considering the collinearity of this factor with language contact ideology, which is a stronger correlation and in the expected direction such that those who are more liberal-identifying hold a more receptive language contact ideology ( $r=.382$ ). This could explain why they fare better at imitating the forms of Indonesia-framed nonce words orthographically represented with foreign-indexed diacritics, relative to their imitation of Iraq-framed nonce words and in comparison to those who are more conservative-identifying. Again, though, this should be treated with caution, as a model replacing the effect of political identity with that of language contact ideology still fares significantly worse at fitting the data ( $p<.0001$ ).

It appears that the social factor significantly explaining the observed variation in the strongest and most directly interpretable way is that of globalist/nationalist alignment: Those who align with a more nationalist ideology are those who show less faithfulness to the exposure forms of new words they've heard when repeating them. This parallels the observations from Chapter 2, where those who are more nationalist-aligning are those who are more likely to pronounce established loanwords in ways that less closely resemble their source forms. This parallel patterning of new and established words provides apparent-time support for the theory that the sociolinguistic variation of loanwords stems from how imitation of the source form at the early stages of a borrowing was mediated by the same social factors.

### **4.3.2 Phonetic imitation**

This section turns to the the subtle phonetic variables manipulated (and more tightly controlled) to analyze for socially mediated phonetic imitation. Before turning to the results, the methods of acoustic analysis for each variable and its imitation are detailed.

### 4.3.2.1 Acoustic analysis and imitation scoring

The variable of GLIDING will be discussed first. This variable was analyzed following Jagers' (2018) documentation of its acoustic character in American English, finding the earliness of transition to the following vowel to be the most consistent cue to the [j]-[i] distinction, with glide-vowel [jə] sequences having an earlier transition to [ə] than hiatus [iə] sequences. Praat acoustic analysis software was used to identify and segment utterances of key pre-existing words from the initial sentence reading task (see Table 21). The vocalic sequence of interest within each word was then identified and segmented. Like Jagers and others (2018; Chitoran 2002; Ren 1986), transition earliness was represented by the measurement of F2 Max Time: the amount of time between the vocalic onset and the F2 maximum, the latest point in time before F2 begins to fall toward the F2 target of following vowel.<sup>6</sup> For each participant, this measurement was averaged across all pre-existing words of each respective variant.<sup>7</sup> As Figure 23 demonstrates, this measurement method does capture the distinction at hand: Participants' average transition in [jə]-expecting pre-existing words is earlier than that in [iə]-expecting pre-existing words. Figure 23 also demonstrates that this distinction was reflected in the stimuli as well, with colored shape points representing the average of each variant per framing speaker.

---

<sup>6</sup>Vocalic onset was identified as the point in time when both F1 and F2 became visible and unattenuated, using Praat's spectrogram view, after the preceding consonant. Vocalic offset was the reverse. For Praat's Gaussian formant tracking, a time step of 0.01s and window length of 0.03s were used. Formant range was individually optimized per speaker, starting with the usual default of 5500Hz for female speakers and 5000Hz for male speakers and adjusting from there. Hand measurement was used for any utterances in which Praat's formant tracker was still visibly failing in spite of individual adjustment.

<sup>7</sup>Some utterances were excluded from analysis: any utterances that clearly did not match the expected production (e.g., an utterance of *hernia* that was confidently perceived as [jə]), and any that sounded like a disfluency or suggested the participant might not have recognized or known the word beforehand (e.g., an utterance of *pneumonia* as [pənəmóɪə]).

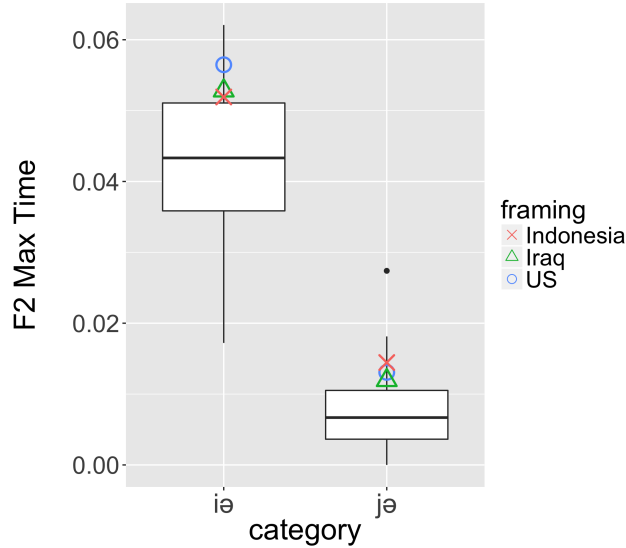


Figure 23: Baseline transition earliness of [jə] vs. [iə]

The amount of time (s) between the vocalic onset and F2 maximum, before F2's fall toward that of [ə]. Box plots represent the average of each participant's baseline utterances. Colored points represent the average per framing speaker: i.e., the stimuli participants heard.

The same analysis was performed for participants' utterances of nonce words manipulated along this variable. The vocalic sequence of interest was segmented and the amount of time between the vocalic onset and F2 Max Time was measured. For each participant, this measurement was averaged across all utterances of each nonce word per exposure/framing. This measurement was then relativized to each participant's baselines, serving as a method of normalization across participants. For each, a proportional score of Glide Likeness was generated as formulated below. A value closer to 1 would mean that the participant's realization of the nonce word was more similar to their baseline realization of a glide-vowel [jə] sequence in pre-existing words; a value closer to 0 would mean that it was more similar to their baseline realization of [iə] hiatus in pre-existing words. This score was then converted to a Heard Likeness measurement to more directly quantify imitation. The Glide Likeness score was subtracted from 1 for tokens where the exposure form was [iə]. Therefore, with HeardLikeness, a higher score means stronger imitation of the exposure form, no matter the exposure form.

$$1. \text{GlideDiff} = | \text{F2MaxTime}_{\text{nonce word}} - \text{F2MaxTime}_{[\text{jə}] \text{baseline}} |$$

2.  $\text{HiatusDiff} = | \text{F2MaxTime}_{\text{nonce word}} - \text{F2MaxTime}_{[\text{i}\text{ə}]\text{ baseline}} |$
3.  $\text{GlideLikeness} = \text{HiatusDiff} / (\text{GlideDiff} + \text{HiatusDiff})$
4.  $\text{HeardLikeness} = 1 - \text{GlideLikeness}$ , iff exposure variant = [iə]

In Figure 24 (lefthand), the Glide Likeness of each nonce word realization across all participants is plotted, categorized by which variant the participant was exposed to when listening to the short story but not yet teasing apart framing or considering social factors of the individual. Results suggest that, in general, imitation of this variable did occur: Overall, when participants heard the [jə] variant used by the speaker during exposure, their realizations of the same nonce word during the sequel reading task more strongly resembled their own realizations of pre-existing words with [jə] than when they heard the [iə] variant during exposure. This is also visualized in a different way (righthand) with the conversion of Glide Likeness to Heard Likeness.

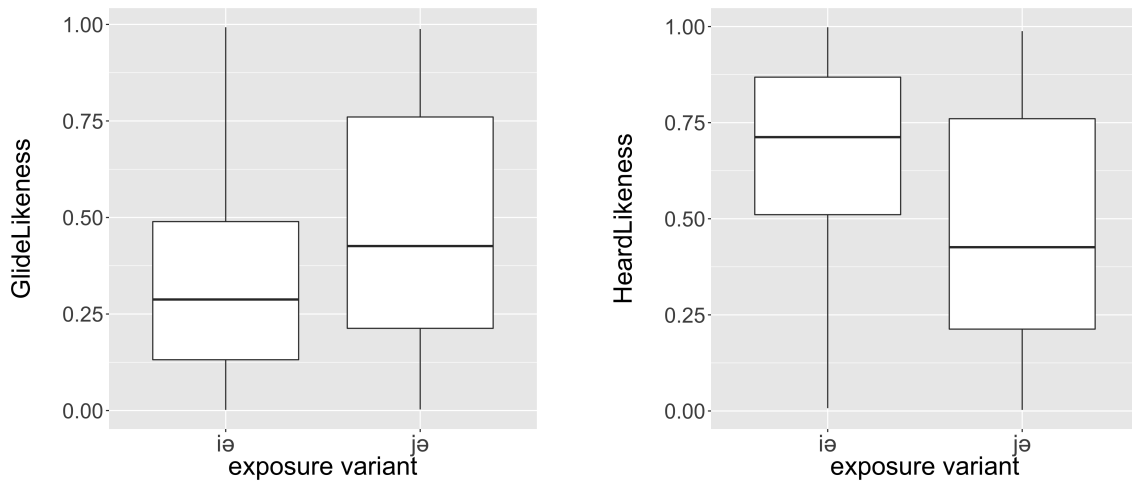


Figure 24: General imitation of exposure form: GLIDING variable

In the lefthand plot, the y-axis represents *GlideLikeness*, the relativized measurement of how a participant's utterances of the nonce word in question resembled their own baseline utterances of pre-existing words with [jə], as compared to [iə]. In the righthand plot, this is converted to *HeardLikeness*, flipping the measurement for cases in which the participant heard the [iə] exposure form.

The fact that there is more overlap of the distributions in Figure 24 than those of the baseline measurements in Figure 23 suggests that there is a high degree of variability in how

strongly participants imitate the exposure form. As Jagers (2018) suggests, while this [j]-[i] distinction is apparent in American English across certain lexemes, it does not yet appear to bear a contrastive load and can be variable for some lexemes (e.g., *piano* [pjáeno]~[piáeno]). This variability was part of the motivation to use the [jə]-vs.-[iə] exposure paradigm, testing for subtle phonetic imitation without the results reaching a near-categorical ceiling or floor. It is also apparent in Figure 24 that there is a preference for hiatus, which Jagers (2018) also observes: While hearing a [jə] exposure form leads participants to use a realization that more strongly resembles their own [jə] form than when they hear [iə] during exposure, the average participant still uses a realization that more strongly resembles their own [iə] form in an absolute sense even after hearing [jə] during exposure. This is apparent by the fact that the mean HeardLikeness score (Figure 24, righthand) is below 0.5 for utterances elicited after the [jə] exposure variant was heard.

The variable of VOWEL REDUCTION and its imitation was calculated similarly to that of gliding. In this case, the F1 and F2 were measured at the midpoint of the vowel of interest for each word of interest from the initial sentence reading task (see Table 21). Each measurement was averaged across utterances pertaining to the respective [ə] and [ɛ] categories to generate a baseline measurement per participant.<sup>8</sup> Figure 25 provides the baseline measurements in the F1×F2 space, averaged and connected for each participant. The distinction is visibly apparent: Overall, participants' [ɛ] realizations are lower than those of [ə]. However, the exact nature of this distinction is different per participant: For some, [ɛ] is has more frontness than [ə] while, for others, [ə] has more frontness; and, for some, the horizontal F2 dimension captures this distinction more so than the vertical F1 dimension. It is for this reason that a measurement of Euclidean Similarity along the F1×F2 dimension was used to analyze imitation, following Babel (2010).

---

<sup>8</sup>Some utterances were excluded from analysis: any utterances that clearly did not match the expected production (e.g., an utterance of *deny* that was confidently perceived as [di...] rather than [də...]). Utterances were also excluded from analysis if the vocoid of interest was entirely deleted.

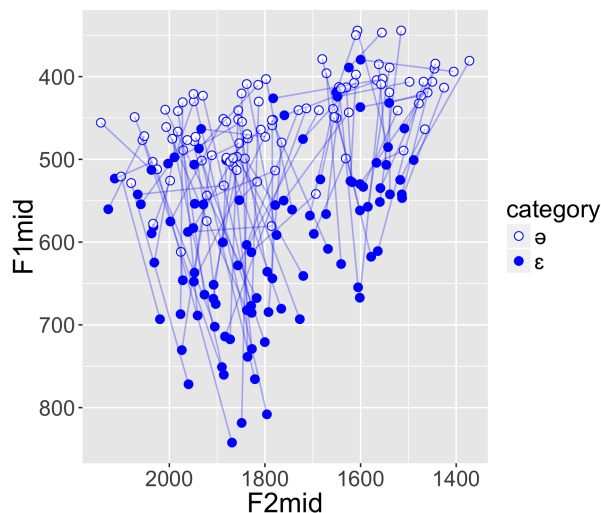


Figure 25: Baseline [ə] vs. [ɛ]

Participants' baseline utterances of the [ə] and [ɛ] categories in the  $F1 \times F2$  dimension, averaged per participant.

For each utterance of a nonce word pertaining to this variable, the F1 and F2 were measured at the midpoint of the vowel of interest. For each participant, this measurement was averaged across all utterances of each nonce word per exposure/framing. (Though utterances were excluded from analysis if the vocoid of interest was entirely deleted.) This measurement was then relativized to each participant's baselines to serve as a method of normalization across participants. For each, a proportional score of Schwa Likeness was generated as formulated below, measuring the Euclidean Distance from each baseline category and the proportional similarity between the two. A Schwa Likeness value closer to 1 would mean that the participant's realization of the nonce word was more similar to their baseline realization of [ə] in pre-existing words; a value closer to 0 would mean that it was more similar to their baseline realization of [ɛ] in pre-existing words. And, like above, this score was then converted to a Heard Likeness measurement to more directly quantify imitation. The Schwa Likeness score was subtracted from 1 for tokens where the exposure form was [ɛ]. Therefore, with HeardLikeness, a higher score means stronger imitation of the exposure form, no matter the exposure form.

1. 
$$\text{EucDist}_{[ə]} = \sqrt{(\text{F1}_{\text{nonce word}} - \text{F1}_{[ə] \text{ baseline}})^2 + (\text{F2}_{\text{nonce word}} - \text{F2}_{[ə] \text{ baseline}})^2}$$

2.  $\text{EucDist}_{[\varepsilon]} = \sqrt{(\text{F1}_{\text{nonce word}} - \text{F1}_{[\varepsilon] \text{ baseline}})^2 + (\text{F2}_{\text{nonce word}} - \text{F2}_{[\varepsilon] \text{ baseline}})^2}$
3.  $\text{SchwaLikeness} = \text{EucDist}_{[\varepsilon]} / (\text{EucDist}_{[\varepsilon]} + \text{EucDist}_{[\emptyset]})$
4.  $\text{HeardLikeness} = 1 - \text{SchwaLikeness}$ , iff exposure variant =  $[\varepsilon]$

An initial survey of the data for this variable suggests that imitation is not robustly observable at the general level, unlike that which was observable for the gliding variable. Figure 26 plots the Schwa Likeness (lefthand), and its conversion to Heard Likeness (righthand). It is apparent that most participants' nonce word pronunciations resembled something more like their own  $[\emptyset]$  categories than  $[\varepsilon]$ . This seems fairly consistent and not much affected by whether the participant heard a pronunciation of  $[\emptyset]$  or  $[\varepsilon]$  during exposure. As is observable in the Heard Likeness view (righthand), a majority of utterances elicited after an  $[\varepsilon]$  still more strongly resembled the participant's  $[\emptyset]$  category (Heard Likeness  $< 0.5$ ). This foreshadows what will be observed below, that no significant imitation is observable for this variable.

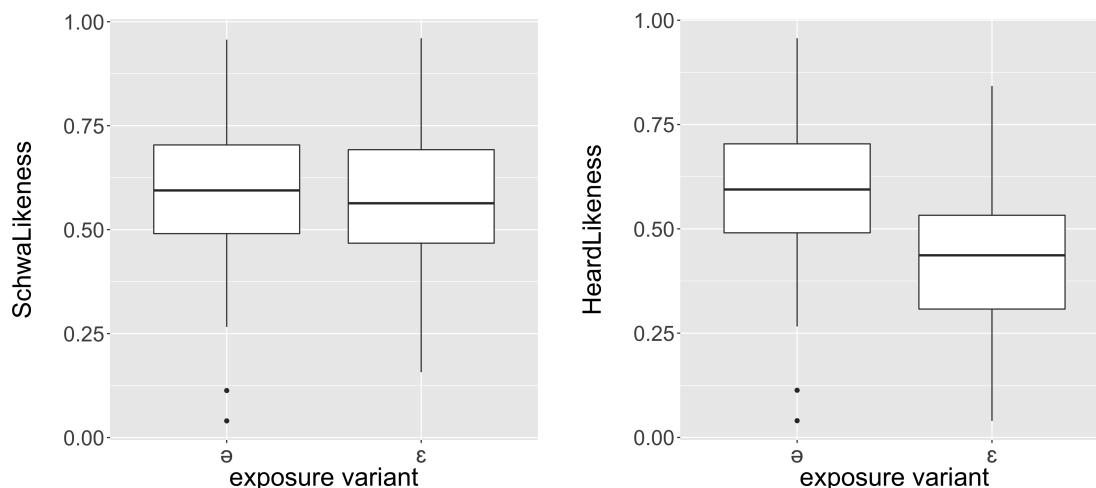


Figure 26: General imitation of exposure form: VOWEL REDUCTION variable  
 In the lefthand plot, the y-axis represents SchwaLikeness, the relativized measurement of how a participant's utterances of the nonce word in question resembled their own baseline utterances of pre-existing words with  $[\emptyset]$ , as compared to  $[\varepsilon]$ . In the righthand plot, this is converted to HeardLikeness, flipping the measurement for cases in which the participant heard the  $[\varepsilon]$  exposure form.

When assessing whether the stimuli adequately exhibited a distinction between the two variants, Figure 27 suggests that these were distinct. However, it appears that this

distinction in the stimuli between the two variants was realized more so along the front/back dimension in comparison to participants' baseline measurements. This could perhaps explain why no imitation of this variable appears to have occurred.

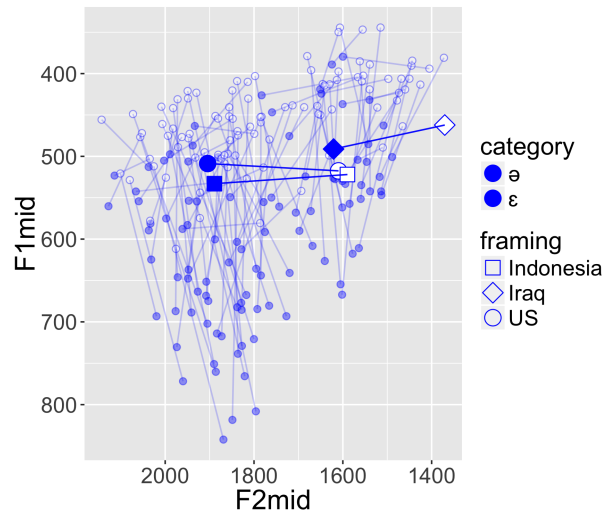


Figure 27: Stimulus [ə] vs. [ɛ] in comparison to participant baselines  
Stimulus speaker realizations of the [ə] and [ɛ] exposure variants in the  $F1 \times F2$  dimension, in comparison to participant baselines.

A potential confound is that the stimuli and participants' non-baseline utterances are both in an unstressed position (hence the variable name “vowel reduction”), while participants' baseline [ɛ] measurements were from utterances of words in which [ɛ] was a stressed vowel (e.g., *dentist* [dɛntɪst], cf. nonce stimulus [dənɪɪ]~[dɛnɪɪ]). When controlled for an unstressed environment, it's possible that participants' utterances also exhibited this distinction (and any imitation of it) along the front/back dimension in line with its realization in the stimuli. However, even when examining whether the exposure variant influenced participants' productions along the front/back dimension without relativizing at all to their baseline measurements, Figure 28 suggests that imitation is still not apparent.



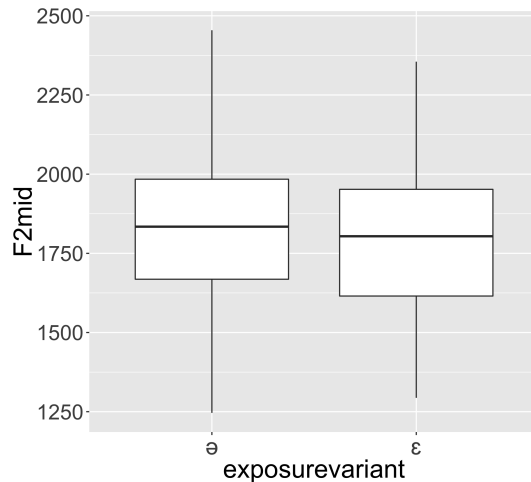


Figure 28: F2 by [ə] vs. [ɛ] exposure: VOWEL REDUCTION variable  
 F2 of participants' utterances of vowel reduction stimuli (non-relativized to baseline measurements), grouped by exposure variant heard.

#### 4.3.2.2 Phonetic imitation results

Of primary interest in this study is how the social factors of the individual might mediate imitation, as well as how they might interact with how the nonce word is framed. For example, we want to know if those more globalist-aligning are those who more strongly preserve the phonetic form of nonce loans and/or if those with more negative attitudes toward Iraq show especially weaker preservation of the phonetic form of nonce loans they consider Iraq-sourced. However, we have also seen that the sound variable itself may mitigate imitation, where each variable seems to have a variant preferred by default and one variable (that of vowel reduction) appears to exhibit little to no imitation.

The difference in imitation behavior across the manipulated variables is also apparent when surveying the data in a preliminary search for how imitation is mediated by social factors of the individual. Figure 29 plots the Heard Likeness imitation score for each participant's nonce word realization by how strongly that participant aligned with a globalist or nationalist ideology, not yet teasing apart framing or exposure variable. Regarding both the variation of established loanwords (Chapter 2) and the imitation of more categorical sound variables in nonce words (analyzed above: §4.3.1), globalist/nationalist alignment was the

strongest predictor of imitation: the use of more or less source-like variants of established loanwords, or the faithfulness to the exposure form of nonce words. Once again this appears to be the strongest predictor, at least at the general level: As is apparent in Figure 29, those who more strongly align with a globalist ideology are those whose utterances of nonce words show a higher Heard Likeness score, meaning stronger imitation of the exposure form. Another way to interpret this would be to consider what the Heard Likeness score means: For those who are more nationalist-aligning, the Heard Likeness score is not significantly different from a value of 0.5, suggesting no effect of the exposure form on subsequent production. Figure 30 provides the analogous plots for each other social factor of interest, showing that the effect is not as robust for them.

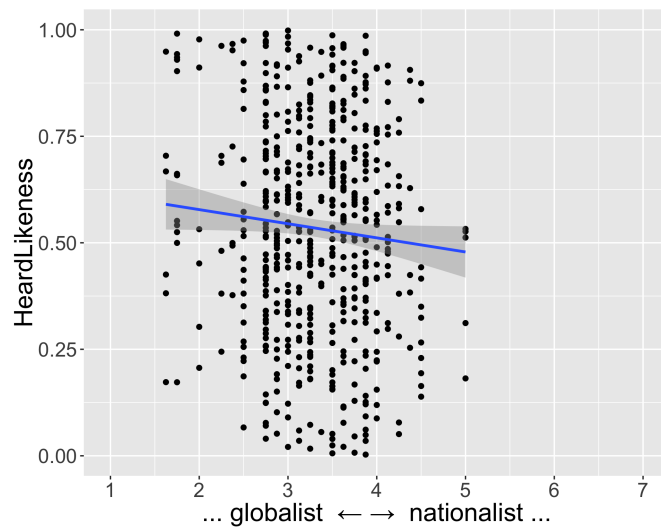


Figure 29: General phonetic imitation: globalist/nationalist alignment  
 HeardLikeness (y-axis) by the participant's globalist/nationalist score based on their responses to the Likert questionnaire (x-axis).

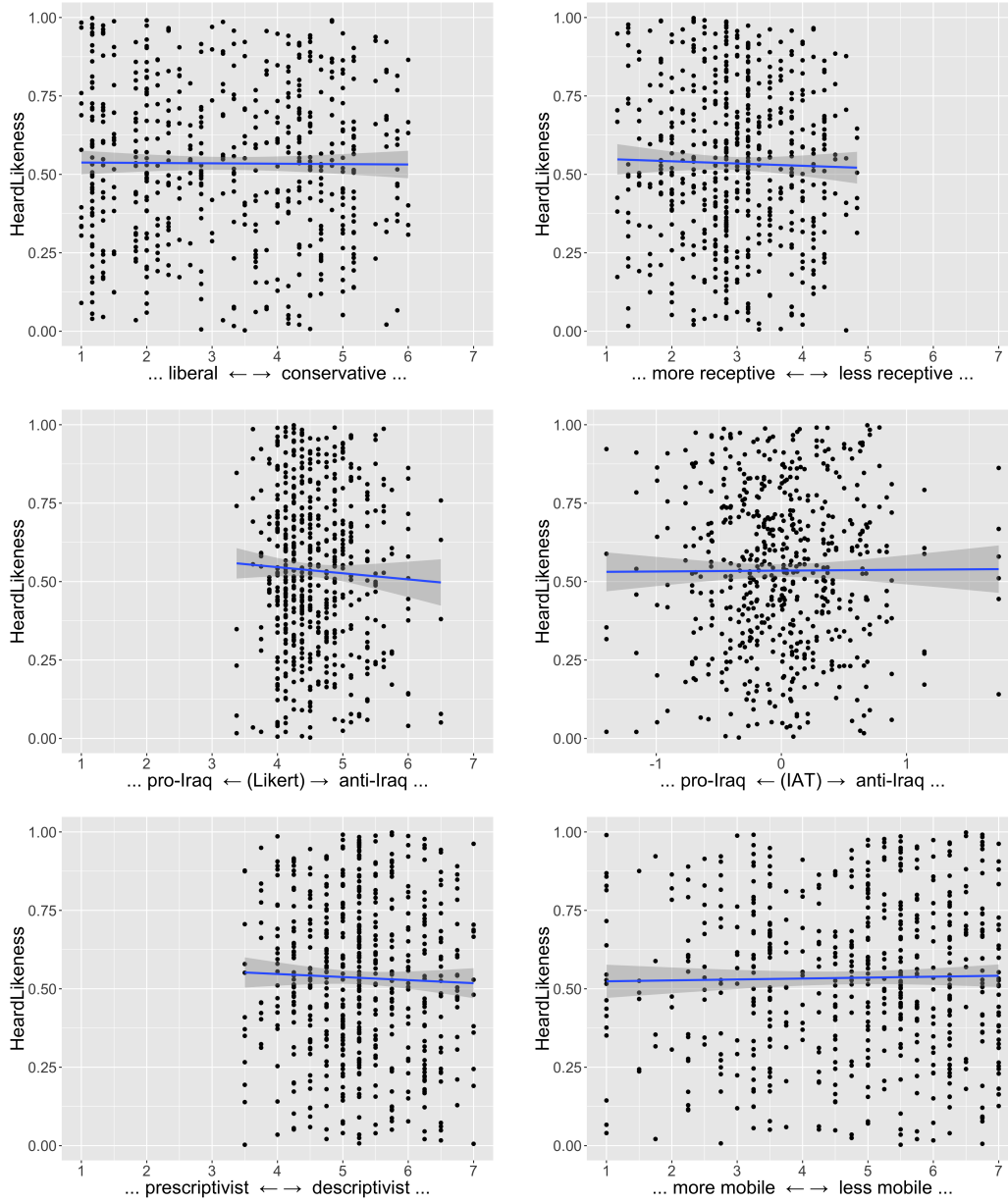


Figure 30: General phonetic imitation: other social factors of the individual

However, as foreshadowed above, Figure 31 suggests that this effect is strongest in (and perhaps completely carried by) the results for the imitation of gliding. For the gliding variable (lefthand), we see that those who are more globalist-aligning show a stronger influence of the exposure form, producing something more [jə]-like when having heard [jə] and vice versa, while the pronunciations by those who are more nationalist-aligning appear not to be influenced by the exposure form they heard. Meanwhile, for the vowel reduction variable

(righthand), even the pronunciations by those who are more globalist-aligning do not appear to have been affected by the exposure form the participant heard.

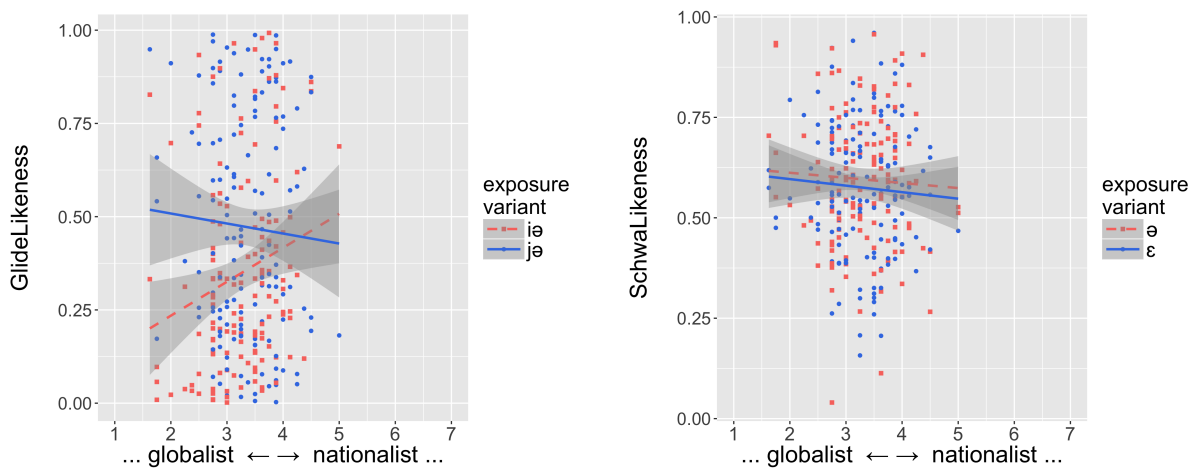


Figure 31: Phonetic imitation: globalist/nationalist alignment, across manipulated variables. In each plot, the x-axis represents the participant’s globalist/nationalist score based on their responses to the Likert questionnaire. The lefthand plot is the data for the gliding variable grouped by exposure variant, with the y-axis representing the GlideLikeness score. The righthand plot is the data for the vowel reduction variable grouped by exposure variant, with the y-axis representing the SchwaLikeness score.

Statistical analysis of the data resembles that performed in Chapter 2 regarding the variation of established loanwords and that regarding the categorical sound imitation of nonce words performed earlier in this chapter (§4.3.1): the generation of mixed-effects models with predictors of interest and a step-up comparison of those models to test for significance of the effects of those predictors. The dependent variable of these models was Heard Likeness, the participant-relativized score of phonetic imitation described above (§4.3.2.1). Seeing that for each variable there appears to be a default/preferred pronunciation, the base model incorporated a random intercept per exposure variant: i.e., A participant who heard the less default exposure form might use a pronunciation that more closely resembles it than had they heard the other exposure form, but in absolute terms it might still more strongly resemble the default pronunciation. Unlike in the categorical imitation examined earlier, the phonological structure and phonography of the target words of each variable was tightly controlled (e.g., the letter <i> was used across all conditions for the gliding variable, and the letter <e> for the vowel reduction variable). A random effect for word was therefore

not motivated (moreover, for each participant, word would conflate with both the factors of framing and exposure variant). Step-up modeling then added individual factors of interest as fixed effects to test for significant improvement of data fit. For each, this included testing whether an interaction term with the factor of framing was motivated.<sup>9</sup>

Confirming the difference between the two variables foreshadowed above, a separate round of model comparisons was performed on the data pertinent to each, which found no social factors of the individual to significantly influence imitation of the vowel reduction variable. The results discussed will therefore exclusively concern participants' imitation of the gliding variable. Table 24. Appendix G provides a more detailed summary of the step-up modeling performed.

FACTOR	INTERACTION (if any)	$\beta$	$\sigma$	$t$	$p$
(intercept)		.561	.083	6.785	
globalist/nationalist alignment		-.036	.017	-2.131	.031
Iraq-directed attitude (IAT)	framing = US	.038	.029	1.344	.025
Iraq-directed attitude (IAT)	framing = Indonesia	.056	.029	1.93	
Iraq-directed attitude (IAT)	framing = Iraq	-.056	.029	-1.94	

Table 24: Step-up comparison results: Phonetic imitation (gliding)

Results of final linear mixed-effects model after step-up comparison of independent variables as predictors. Each factor's  $p$ -value comes from a Chi-square ANOVA test between the full model and a model excluding the component of interest. ("Intercept" = US framing.)

As foreshadowed in the comparison of Figures 29 and 30, the factor of one's globalist/nationalist alignment is the strongest general predictor of imitation, with statistical results not motivating the inclusion of an interaction with framing. Those who are more globalist-aligning are those whose pronunciations of nonce loanwords, as well as nonce words framed as non-loans, more strongly resemble the sound form they heard disseminated to them.

One additional factor was also found to exhibit a significant effect. As demonstrated in Figure 32, a participant's IAT score significantly influenced their imitation, depending on

<sup>9</sup>And, like above, in the following statistical reporting, effect coefficients ( $\beta$ ), standard deviations ( $\sigma$ ), and  $t$ -values ( $t$ ) come from the final model including all fixed effects identified as significant contributors to model fit;  $p$ -values ( $p$ ) come from a Chi-square ANOVA test between that full model and a model excluding the effect of interest (Winter 2014).

the framing: Those with stronger anti-Iraq biases as measured by the IAT showed weaker imitation of the exposure forms of nonce words framed to them as Iraq-sourced loans relative to their imitation of other nonce words.

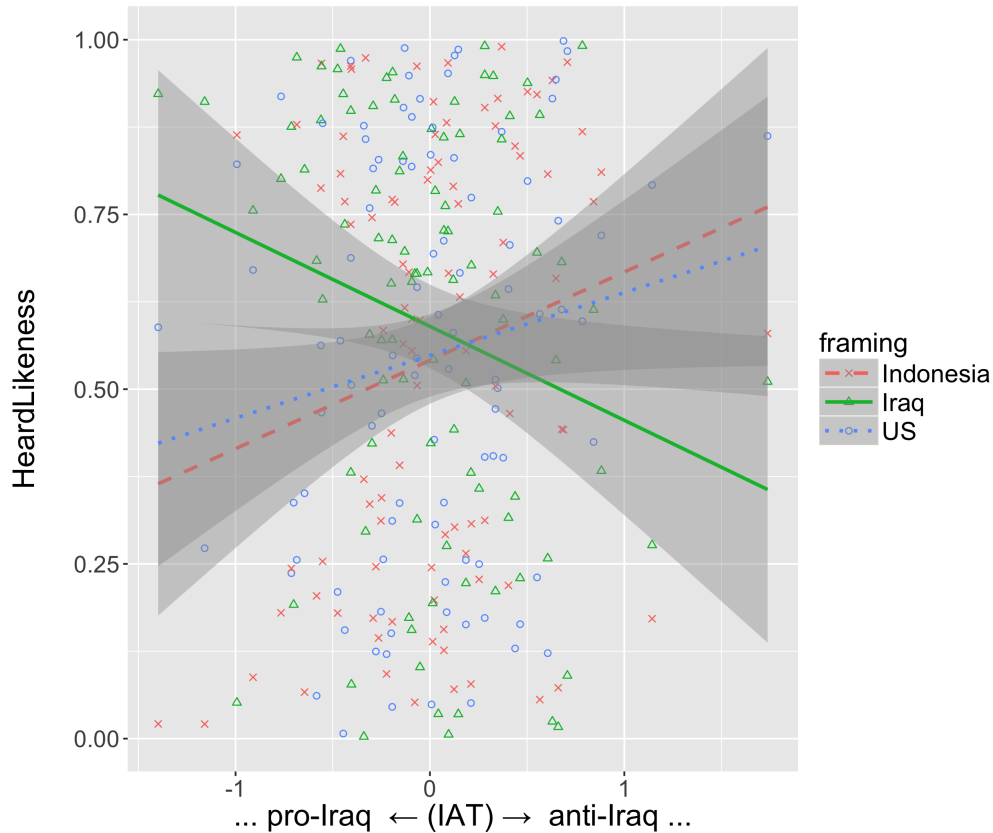


Figure 32: Phonetic imitation of gliding: place-directed IAT  
 HeardLikeness (y-axis) by the participant's score on the Implicit Association Test between Iraq and Indonesia (x-axis).

In summary, we see the same pattern observed above (§4.3.1): Those who align with a more nationalist ideology are those who show less faithfulness to the exposure forms of new words they've heard when repeating them, at least along the one variable (gliding) that did appear to exhibit any effect suggesting imitation of the exposure form. This once again parallels the observations from Chapter 2, where those who are more nationalist-aligning are those who are more likely to pronounce established loanwords in ways that less closely resemble their source forms. This parallel patterning of new and established words provides apparent-time support for the theory that the sociolinguistic variation of loanwords stems

from how imitation of the source form at the early stages of a borrowing was socially mediated by the same social factors. We also see that attitude toward the (posited/framed) source of the loanword can have an effect, empirically supporting Weinreich's (1968, 27) hypothesis. Those who hold more negative implicit biases regarding Iraq show less faithfulness to the exposure form of a loanword they consider Iraq-sourced.

#### 4.4 Discussion and Conclusions

It appears that the source form replication of new, incoming loanwords is socially mediated in ways similar to the variation of established loanwords. Those who are more likely to use more source-like variants of established loanwords are also those who more strongly imitate the sound forms of new loanwords they hear and further disseminate. The strongest and most consistent social mediator appears to be one's alignment with a globalist or nationalist ideology. As observed in Chapter 2, this is the strongest predictor of the variation of established loanwords between more and less source-like variants, where those who are more globalist-aligning are those who are more likely to use more source-like variants. This also appears to hold consistently across loanwords from various sources. A parallel effect is observed in this chapter's study of the sound form replication of new loanwords. Those who are more globalist-aligning are those who pronounce new loanwords in ways more faithful to the sound forms they heard. This is also consistent across loanwords framed as Indonesia-sourced and Iraq-sourced.

In an apparent-time way, this suggests that sociolinguistic loanword variation is a result of loanword adaptation being mediated by the same social factors that condition the variation of established loanwords. That is, when we see older loans and younger loans patterning in parallel with respect to source form replication, we might infer that the observable variation of older loans between more and less source-like pronunciations results from social mediation of the degree to which they were imitated along the path of adaptation from the very beginning, just like the imitation of their younger counterpart loans appears to

be. In this study, when choosing between a more [jə]-like or [iə]-like pronunciation, more nationalist-aligning speakers show relatively little influence of whether they heard a more [jə]-like or [iə]-like pronunciation beforehand. More nationalist-aligning speakers are also those who are more likely to use an [æ] pronunciation of *Iraq* even though an [ɑ] pronunciation would be more source-like. We might, therefore, imagine that during *Iraq*'s path of establishment, it was those who were more nationalist-aligning who might have been less influenced by what they heard when choosing between [ɑ] and [æ] pronunciations, therefore being those more likely to use an [æ] pronunciation and give it more weight as a candidate for widespread establishment.

In fact, this effect of globalist/nationalist alignment as a mediator of imitation is surprisingly consistent, seeing that the same pattern also holds for nonce words framed as non-loans (i.e., those framed as unfamiliar English words in a US setting). In the design of this study, it was anticipated that there might be an interaction of this effect with loanword status: that those who are more globalist-aligning would show a stronger imitation of nonce loans relative to non-loans in comparison with those who are more nationalist-aligning. Instead, they show stronger imitation across the board without an apparent effect of loanword status.

Of course, this lack of an interaction with loanword status does not rule out the conclusion that loanword variation stems from socially mediated adaptation: If those who are more globalist-aligning are better sound form imitators of new loanwords it makes sense why they are also more likely to use more source-like pronunciations of established loanwords, regardless of whether they also show stronger imitation of non-loans. This observation could even further enlighten our understanding of the socially mediated path of loanword adaptation and establishment. Continuing the application of these findings to the adaptation of *Iraq*, this could explain the path beyond the initial stage of introduction. At the initial stage, those who are more nationalist-aligning may show a weaker influence of the foreign source form [ʕi'raq] in choosing between [ɑ] and [æ] pronunciations. But, say a speaker never



heard the source form to begin with. The findings in this study suggest that, even when hearing a new word uttered by someone they consider non-foreign and a native speaker of the borrowing language, someone who is more nationalist-aligning may still be less influenced by the form they hear when pronouncing the word in the future.

This broad effect could reflect something even more core to one's globalist/nationalist alignment. As mentioned at the outset of this study (§4.1.1), Yu and colleagues (2013) observe that one's phonetic imitation of another speaker's VOT is mediated by their attitude regarding that speaker, where a more positive attitude leads to stronger imitation. In addition, Yu and colleagues observe that imitation is influenced by a participant's score of Openness, one of the personality facets measured in the Big Five Inventory (John et al. 1991; John et al. 2008) used in social psychological research and diagnostics. This personality facet represents one's degree of interest and level of comfort in new experiences, resembling globalism in an even broader and more abstract sense. And, like globalism, a higher Openness score has also been observed to correlate with a more liberal political identity (Caprara et al. 2003; 2006; Chirumbolo and Leone 2010; Jost et al. 2003).

This does, however, open the door for future questions to be examined. As observed by Yu and colleagues (2013) regarding Openness, it is possible that those who are more nationalist-aligning exhibit less accommodation to fellow speakers of their native language, even without any indexation of the speaker or the words they're using as foreign. The effect of foreignness may still be at play, though, and should not yet be completely ruled out. A further study might concurrently measure both the participant's globalist/nationalist alignment and one's personality facet of Openness as predictors of imitation with special attention to teasing them apart and comparing them as predictors, much like this dissertation's aim of teasing apart and comparing political identity from other relevant social factors. Combined with a manipulation of loanword status, it may still be observed that the influence of globalist/nationalist alignment is stronger in the imitation of loanwords than the influence of Openness, which may be more general.

The effect of source-directed attitude also resembles that observed in Chapter 2, albeit in a complicated way. Recall that those with more positive attitudes regarding Arabs and the Middle East were more likely to use the more source-like pronunciations of *Iran* and *Iraq*. However, this effect was more observable when measuring source-directed attitude explicitly via the Likert questionnaire rather than implicitly via the Implicit Association Test. Here however, with the imitation of new loanwords, it is one's implicit biases regarding Iraq that better predict preservation of the source form. This resembles Babel's (2010) observation that New Zealanders' phonetic imitation of an Australian speaker is better predicted by their implicit biases regarding Australians than by the explicit manipulation of their attitude toward the speaker.

These findings may suggest that implicitly measured attitude can be a predictor of imitation on a short-term basis as a more immediate, online effect. But, in terms of what a speaker converges on in a more long-term sense, especially for words in actual use in the community, this effect may be overridden by one's explicit attitudes. One with a more negative implicit attitude regarding Iraq might be less inclined to faithfully replicate the pronunciation of an Iraq-sourced loanword they initially hear and disseminate. And, this implicit attitude correlates with explicit attitude (Nosek et al. 2007), explaining how those with more negative attitudes that they do explicitly espouse also show a weaker replication of the source form for words that have become more established and widespread. But, implicit and explicit attitudes are not necessarily one and the same. Meta-analyses of social psychological research suggest that implicit bias measures alone can be rather inconsistent in predicting biased behavior (Oswald et al. 2013), that they are better predictors of behavior when they are more strongly aligned with explicit self-reported measures (Greenwald et al. 2009), and that changes in implicitly measured bias do not neatly lead to changes in biased behavior (Forscher et al. 2018 [preprint]). (And, as Nosek and colleagues [2007] also observe, attitudes regarding Arabs, the Middle East, and Islam tend to be stronger explicitly than implicitly, unlike other group-directed biases.) As a word becomes established with a more

frequent and widespread usage, the short-term, immediate effect of implicit bias may therefore be overridden. One who has a negative implicit attitude regarding Iraq but who does not as strongly align with this attitude explicitly may be more likely to adjust their speech in alignment with their explicit attitude and converge with those using a more source-like pronunciation. Whether implicit or explicit, though, these findings provide empirical support to Weinreich’s (1968, 27) hypothesis that one’s attitude toward the source mediates their preservation of the source form of a loanword.

Another complication in the results of this study is the observation that imitation was really only observed and socially mediated for the [jə]~[iə] gliding variable and not for the [ə]~[ɛ] vowel reduction variable. It may be notable that most previous research analyzing phonetic imitation has focused on imitation in stressed syllables: Babel’s (2010) study of vowel quality imitation only examines that in monosyllabic words, and Yu and colleagues’ (2013) examine the VOT imitation of aspirated voiceless stops in stressed, word-initial syllables. It may be that the vowel reduction variable was less salient to participants due to its location in an unstressed syllable ([Cə'CVC]~[Cɛ'CVC]). However, it is notable that the gliding variable was also located in an unstressed syllable ([C'VVCjə]~[C'VCiə]), and Babel also observes that even in the imitation of stressed vowels some categories exhibit stronger phonetic imitation effects than others. It could be that prosody overrode phonetic imitation for this variable, though, given that it was even more directly connected to prosody: i.e., the preservation of a [ɛ] realization in the unstressed syllable of interest here directly conflicted with the English prosodic rule favoring that [ɛ] be reduced to [ə] in such an environment.

The gliding variable, on the other hand, could be considered something in between a phonetic and categorical variable. Jagers (2018) notes that there appears to be a “distinction” between [j] and [i] in American English, but that this may not be considered a “contrast”: There is no evidence of exact minimally paired words in which it bears a contrastive load, and there is observable variability within words across speakers (e.g., *piano*

as [pjáeno]~[piáeno]). Rather than this variable conflating with prosody, either variant is possible in the same phonological environment and a speaker is free to choose one. As seems to be the case here, this choice is more free and random for some speakers, while others' choice between the two options is more strongly influenced by what they heard used during exposure.

As was observed before analyzing the imitation of gliding and vowel reduction however (§4.3.1), even the imitation of more categorical sounds and features can be subject to variability. Especially for nonce words framed as loanwords, some participants' utterances exhibited substitutions, deletions, metatheses, and stress shifts, which all could be unquestionably contrastive in words native to American English. This, too, is observed to pattern in the same way, with more nationalist-aligning participants being those more likely to exhibit such divergences from the exposure form. A note of caution was made, though, that much of this divergence (though not all) might be attributable to phonographic bias (i.e., the preference for associating certain orthographic representations with certain sounds). Therefore, this socially mediated imitation (termed as 'categorical imitation' above (§4.3.1) might best be considered socially mediated suppression of phonographic bias in the preservation of the sound form.

This motivates a further question for future research: Are there certain orthographies or phonographic associations that might be indexed as especially native or foreign? Boberg's (1997; 1999) analyses suggest that American English speakers index an association of /æ/ with an orthographic representation of <a> as more native and less foreign than /ɑ/. Janda et al. (1994) further suggest that this indexation might be influenced by the relative dynamics of language contact. For example, they observe that the French <j> → /ʒ/ spelling-to-sound mapping appears to be overgeneralized and applied to foreign words as a broad class, even in cases where /ɟ/—also a readily available sound in the American English inventory—would be considered a more accurate replication of the source form (e.g., *Beijing*, *Taj Mahal*). Further experimentation could test the roots and productivity of such 'hyperforeignisms', as

Janda et al. refer to them. For example, experimentation could test how strongly participants apply this <j> → /ʒ/ spelling-to-sound mapping to new words, whether this is stronger for new words framed as loanwords, and whether this is mediated by how globalist-aligning they are or, say, the degree of prestige they ascribe to French.

A related (or perhaps precursor) question, however, is whether this mediated faithfulness to what participants heard (and its possible overriding of phonographic biases) was actually faithfulness to what they remembered hearing at all. In other words, did more nationalist-aligning participants show weaker imitation because they had a weaker inclination to replicate the exposure form? Or, did they show weaker imitation because they were less attuned to and/or motivated to store the phonetic detail of what they heard in the first place? Discussion in Chapter 5 will return to this question and further elaborate on considerations of sound processing and memory as they relate to imitation, as well as motivating further research to test which subprocess of imitation (perception/storage vs. production) is socially mediated.

This study provides an important complement to the observations of Chapters 2 and 3 of this dissertation, with the aim of deepening our understanding of loanword variation and adaptation. In Chapter 2, we observed the sociolinguistic variation of established loanwords between more and less source-like pronunciations. This was observable beyond just the variation of Arab-sourced *Iran* and *Iraq* that had previously been observed (Hall-Lew et al. 2010; Silva et al. 2011; Hall-Lew et al. 2012), with loanwords as a broad class largely patterning together. This variation was best explained by the similarly broad factor of participants' globalist/nationalist alignment, rather than political identity. It was also observable along participants' source directed attitude, more so that which they were explicitly willing to espouse rather than that which was measured implicitly. It appears to have gained a political indexicality in its own right, though, while the variation along political identity is mediated by how politically charged the loanword's source is. In Chapter 3, we observed that this variation is similarly indexed upon perception, with the use of more source-like loanword

pronunciations indexing one as more global and seemingly gaining them prestige in what we might consider to be a global linguistic market. A political indexicality is also activated upon perception, though in a similarly second-order fashion: only when participants think explicitly and metalinguistically about the variable of interest. The observations in *this* chapter help shed light on the roots of how this variation might have arisen in the first place. We see that those who are more likely to use more source-like pronunciations of established words are those who more strongly replicate the source forms of new, incoming loanwords in the first place. These speakers are, once again, those who are more globalist-aligning. We also see an influence of one's attitude toward the particular source of the loanword. However, at this beginning stage of the adaptation process, it is one's implicit attitude toward the source that mediates their replication of the source form. As we might infer by relating this to the observations in Chapter 2, one's explicit attitudes take over in governing how source-like a pronunciation they use as a loanword becomes more established. And, as observed in Chapters 2 and 3, the seemingly strongest, most core predictor of source form preservation (globalist/nationalist alignment) remains the strongest, first-order indexicality of loanword variation, while also picking up related indexicalities regarding both how it patterns and how it is perceived.

## Chapter 5

### Conclusion

#### 5.1 Summary of findings

In Chapter 1 of this dissertation, the borrowing, adaptation, and variation of loanwords is introduced as a topic of interest to the linguistic sound system, sociolinguistics, and language contact. Regarding the sound system, loanword adaptation and possible differences between languages in how loanwords are adapted may tell us how sounds or sound sequences (especially those that are unfamiliar) are perceived, stored, and/or (re)produced. But, sometimes there is variation still observable *within* a language between pronunciations that more or less closely resemble a loanword's source form, in spite of both being available in the borrowing language's sound system. Such cases may be examined to shine light on the social relation between the speaker and the source, rather than the similarities or discrepancies between the sound systems of the borrowing language and the source language. Furthermore, it is considered that the social relation of interest might not only be the singular language contact situation particular to a loanword's specific source language. Instead, loanwords as a broad class may pattern together, similarly reflecting something about the social relation of the speaker with respect to foreign languages as a collective group. This dissertation therefore examines the variation of loanwords in American English between more and less source-like pronunciations, especially considering how loanwords of diverse sources and sound variables may pattern together, reflecting the way speakers align themselves with respect to languages (and people who speak them) which they consider foreign.

The first study in this dissertation (Chapter 2) sets the stage by examining variation in the pronunciation of established loanwords by American English speakers. First, a diverse while interconnected suite of social factors is identified as potential predictors of this varia-

tion, set against a backdrop that is Hall-Lew et al.'s (2010) previous study of the variation of *Iraq* in American English between more source-like [aɪˈræk] and less source-like [ɪˈrɑk] pronunciations, which they observed to pattern with US politicians' political identity. Multiple scholarly studies regarding language contact, loanword variation, and loanword adaptation are considered in motivating the consideration of other potential predictors, complemented by qualitative analysis of how loanword variation appears to be indexed and evaluated in public discourse. Source-directed attitude, language contact ideology, and globalist/nationalist ideology are proposed as more direct predictors, motivating the hypothesis that political loanword variation may fall out from and therefore reflect their correlation with political identity.

In a series of experiments eliciting speakers' utterances of variable established loanwords, these social factors are compared alongside that of political identity as predictors of loanword variation. Observations suggest that this variation is not as directly attributable to political identity as Hall-Lew et al.'s findings and interpretations might suggest. Instead, it is observed that speakers' globalist/nationalist alignment is the strongest predictor when accounted for alongside political identity, with those who are more globalist-aligning being those who are more likely to use more source-like pronunciations, across a diverse array of loanwords from various sources. It is argued that political variation is therefore second-order (Silverstein 2003) to that of globalist/nationalist alignment: that political variation can be explained in large part as a result of political identity's correlation with globalist/nationalist alignment (especially in the US political zeitgeist). However, loanword variation is still observable along political identity, at least for words that might be considered more politically charged: those related to Arabs, Islam, and Latin America. This observation is interpreted as a sign that the second-order political indexation is gaining strength (at least for words of more charged sources or associations) and not solely a byproduct of the correlation of political identity with the more primary predictor that is globalist/nationalist alignment.

The next study (Chapter 3) examines how American English speakers perceive and



think about loanword variation, as well as how perceptions of it line up with the way it patterns in actual usage. It is observed that the perceptual indexation in large part resembles the way loanword variation patterns in production. The use of more source-like pronunciations is associated with global orientation, as well as prestige and linguistic security. The nexus of these indexations is interpreted to suggest that the use of more source-like pronunciations of loanwords can be a resource for capital in what might be considered the global linguistic market (Bourdieu 1977; Piller 2001; Zhang 2005; Blommaert 2010). Additionally, political indexation is once again apparent but in a second-order fashion: Participants do not index loanword variation with political identity in an implicit, matched-guise test; however, they do suggest a political indexation in a metalinguistic questionnaire asking explicitly for their opinions about what loanword variation means. Results also suggest that the use of more source-like pronunciations is an in-group prestige marker of sorts, with fellow liberals and self-reported users of more source-like pronunciations, themselves, more strongly considering this usage to index global orientation, liberal political identity, and linguistic security.

The final study (Chapter 4) cycles back to the source, examining how this sociolinguistic variation of loanwords may arise in the first place. Participants are presented with nonce words framed as loanwords and subsequently prompted to utter them after having heard them. These subsequent utterances are examined for how strongly they resemble the exposure form the participant first heard uttered. It appears that those who are more likely to use more source-like variants of established loanwords are also those who more strongly replicate the exposure forms of new loanwords they hear. Once again, globalist alignment is the strongest predictor of source form replication. This suggests that loanword variation may come about by the initial stages of loanword adaptation being socially mediated by social factors similar to those predicting the variation of well-established loanwords. Results also provide direct empirical support for Weinreich's (1968, 27) hypothesis that attitude toward the source of a loanword can mediate imitation. When a nonce loanword is framed as being sourced from Iraq, in particular, those with stronger implicit biases against Iraq show weaker

imitation of its source form.

## **5.2 Theoretical implications**

When put together, these studies further enlighten our understanding of loanword adaptation, providing evidence that initial adaptation can be socially mediated and may be the root of the socially meaningful variation of even well-established loanwords between more and less source-like pronunciations. Furthermore, these observations encourage future research to consider the variation of loanwords as possibly the pattern of a broad class, rather than a pattern specific to a particular loanword or set of loanwords from one particular source. These studies also provide insight regarding indexicality, both the particular indexicality of loanword variation in American English while relevant more generally to the theoretical notions of indexical order (Silverstein 2003) and indexical fields (Eckert 2008). It appears that speakers' globalist/nationalist alignment is what might currently be considered the most core predictor and primary indexation of the use of more source-like pronunciations, amongst predictors considered and tested here. But, the connection of this alignment with US political identity (both by sociological correlation and ideological association) has seemingly led loanword variation to gain a political indexation, seeing that it is especially apparent when speakers think about loanword variation directly and comment on its meaning explicitly. Together, these findings also enlighten our understanding of the relation of English to the global linguistic market. While English is highly dominant in this market, the usage of and faithfulness to loanwords in English suggest that even speakers of this language may not be considered to carry the maximum possible linguistic capital in the global linguistic market simply by speaking English natively.

### **5.2.1 The socially mediated path of adaptation**

A primary contribution of this dissertation is the concurrent analysis of the variation of established loanwords with the adaptation of new loanwords. The findings suggest that

these phenomena are connected: The social factors that predict whether one will use more source-like variants of loanwords that are well-established but nonetheless variable are also those that predict whether one more closely replicates the forms of new loanwords they've heard when subsequently reproducing them.

The apparent-time approach to language variation and change (Labov 1963; Bailey 2002) examines the language usage of younger and older generations of speakers in a community. When there is a significant difference, it is generally (while with some caution) interpreted that this reflects a change in the language of that community: that the present-day speech of older speakers likely doesn't differ too much from the way they spoke years before when they were the age of the present-day younger speakers concurrently examined; therefore, this is considered to reflect a difference between speakers who would otherwise be the same age but measured at different points in real time. That is, the speech of younger and older speakers is examined for differences in linguistic form that can be inferred to mirror differences between speakers of the same age but at different points in time.

This study might be considered an apparent-time approach to loanword variation and adaptation. Except here, younger and older *linguistic forms* (loanwords) are examined for *parallels* in how they pattern across speakers to infer how older linguistic forms may have patterned when they were younger at an earlier point in real time. This application of apparent time to the adaptation of loanwords is not new to the study of loanword phonology. As discussed in Chapter 1, Itô and Mester (1999; see also Saciuk 1969; Paradis and LaCharité 1997) analyze synchronic differences in the phonological adaptations of loanwords in Japanese traceable to different epochs of language contact. They observe that each epoch's borrowings carried with them at least some significant changes to the Japanese phonology, allowing faithful replication of sounds or phonotactic structures not previously allowed, which continued to apply to borrowings from later epochs but not retroactively to earlier borrowings that had already become well-established.

The application of apparent time to the adaptation of loanwords is also not new to the

study of sociolinguistic loanword variation. Poplack et al. (1988) study the synchronic variation of English-to-French loanwords. They observe that older loanwords are more phonologically integrated (i.e., pronounced with fewer sounds maintaining their English realization) than younger loanwords, by which they infer that this path is gradual rather than instantaneous. That is, there appears to be a change, *in form*, across time. However, while the path may be gradual and show a change (*in form*) across time in some cases, the current study suggests that the same social influences may mediate source form replication along all stages of this path. That is, there might not be that much of a change across time in terms of social influences and resultant sociolinguistic patterning.

In this case, regarding loanwords in American English, a consistent effect of speakers' globalist/nationalist alignment is observed across younger and older loanwords. Those who are more globalist-aligning are more likely to use more source-like variants of established loanwords, and they are also those who more closely replicate the pronunciation of a new loanword they encounter. We might therefore infer that, for the established-while-variable older loanwords, it was also more globalist-aligning speakers who more closely replicated the loanwords' source forms at the early stages of their paths of adaptation, allowing for the more source-like pronunciations to remain in use and become established. And, conversely, it may have been more nationalist-aligning speakers who showed more variation and less of an influence of the source form when pronouncing a younger loanword, therefore further disseminating less source-like pronunciations (for reasons like orthographic influence, for example) and giving them weight to also become established variants.

It is also observed that this effect of globalist/nationalist alignment holds for new words framed as non-loans as well. As discussed more thoroughly in the conclusion of Chapter 4, this too may be in harmony with the idea of a socially mediated path of adaptation. Even if a more nationalist-aligning speaker does not hear a new loanword and introduce this variation at the word's initial incorporation into the borrowing language but, instead, hears it further down the dissemination chain, they may still exhibit less faithfulness to the form they hear

and potentially diverge from that form in their own further dissemination. That is, no matter how strongly a word is indexed as foreign, a more nationalist-aligning speaker may still show less influence of the exposure form when subsequently reproducing it.

The above description of the path of adaptation comes with a caveat that these effects should most likely be interpreted at the aggregate level. Loanword adaptation and establishment is a developing process rather than some one-time instance that cements the resultant form or that repeats anew with every speaker or every utterance (Haugen 1950; 1953; Kang 2010; de Jong and Cho 2012). As discussed more thoroughly in the conclusion of Chapter 2, some speakers may simply inherit a pronunciation from their surrounding community of speakers. In some cases, this may be in accord with the observable sociolinguistic variation, where a speaker inherits both a certain pronunciation and identities/ideologies that pattern with it. However, there may also be discord, where a speaker's identity or ideology changes during their lifespan, or where a speaker identifies one way while having inherited the unpredicted variant. But, such a speaker may still behave in accord regarding their treatment of a new word, and/or they may adjust their pronunciation of an established loanword later in life.

An additional caveat is that many factors can play a role in shaping the path of a loanword's adaptation. As reviewed in Chapter 1 (§1.2.2), the similarities or differences between the sound systems of the source and borrowing languages most certainly play a role, as well as other factors like the borrowing language's phonographic biases and the degree to which a loanword has become established and widespread in its use by speakers less and less familiar with the source language. Therefore, it might not be the case that every single sound or loanword is equally likely to result in multiple established variants due to socially induced noise along that path. Many of these other potential influences were controlled in the study undertaken in Chapter 4 with respect to the nonce loanwords examined for phonetic imitation: degree of establishment, licitness in the borrowing language's sound system, consistent orthography within framings, etc. Even then, only one of the two sound variables

tested (gliding, but not vowel reduction) showed a socially mediated effect that could be predicted to result in differing established forms further down the path of adaptation and establishment. As discussed in the conclusion of Chapter 4, this sound variable might have been especially prime for social mediation, given that [jə] vs. [iə] may be phonologically distinct in American English while not necessarily contrastive (Jagers 2018) and that this distinction can be represented by the same orthography without being strongly overridden by a phonographic bias. With such ripe potential for variation, social factors could more easily sway adaptation one way or the other and increase the likelihood for the path to split. However, this does not necessarily mean that all other factors must be perfectly set so that variable adaptation exhibiting the effects of social influences can then and only then be a possibility. Social influences are but additional members of the set of influences, which could very well interact with or suppress the power of other influences such as the sound system or orthography (Thomason and Kaufman 1988; Thomason 2007).

Beyond the effect of globalist/nationalist alignment, we also see, in line with Weinreich's (1968, 27) hypothesis, that attitude toward the particular source can mediate loanword adaptation. In the study of nonce loanwords (Chapter 4), it appears that a speaker with stronger anti-Iraq implicit biases (as measured by the Implicit Association Test [Greenwald et al. 1998]) shows weaker imitation of the exposure form of new loanwords they consider Iraq-sourced. However, in this case, there appears to be a subtle change between old and new loanwords regarding what attitudinal factor best predicts their variation. In the variation of Arab-sourced *Iran* and *Iraq* (Chapter 2), speaker's *explicit* attitudes regarding the Middle East are a stronger predictor, with those expressing more anti-Middle East sentiments being those who are more likely to use less source-like variants.

As discussed in the conclusion of Chapter 4 (and further below: §5.3.2), this could suggest a difference in short-term processing vs. how one settles on a pronunciation in a longer-term way. As Babel (2010) observes, implicit biases regarding a group (also measured by IAT) are a better predictor than more explicitly manipulated attitude of one's phonetic

imitation of a speaker from that group during an online imitation task. In social psychology, however, the exact correlation between measured implicit bias and observable discriminatory behavior remains unclear and much debated (Oswald et al. 2013; Greenwald et al. 2009; Forscher et al. 2018 [preprint]). The current observations suggest that explicit attitudes may take over regarding less online and more long-term behavior (at least linguistic behavior). Speakers with stronger anti-Iraq implicit biases may show less faithfulness to the source forms of new Iraq-sourced loanwords they hear and reproduce soon after; but, as a word becomes more established, even one with such implicit biases may still gravitate toward a more source-like pronunciation of it, in accord with the attitudes they explicitly identify with.

### 5.2.2 Indexicality

The studies in this dissertation compared multiple social factors as predictors of loanword variation (Chapter 2), adaptation (Chapter 4), and perceptual indexation and evaluation (Chapter 3). Each study, in a different way, suggests that the more source-like pronunciation of loanwords is most primarily indexed with globalist alignment. In loanword variation and adaptation, this is the strongest predictor of whether one will use more source-like pronunciations. And, in sociolinguistic perception, a speaker using more source-like pronunciations is significantly indexed (across implicit and explicit methods) as more globally oriented.

A large motivation behind these studies, however, was to examine the complexity and potentially layered nature of the indexicality of loanword variation. Hall-Lew et al. (2010) observe the variation of *Iraq* between more source-like [ɑ] and less source-like [æ] pronunciations to pattern with US political partisanship, with interpretations further suggesting that this may be a linguistic ‘act of [political] identity’ (Le Page and Tabouret-Keller 1985). This dissertation’s research program was motivated to dissect Hall-Lew et al.’s observations and interpretations, hypothesizing that this variation may not be so directly and strongly connected to political identity. Instead, this research pursued the analysis of loanword variation

along political identity while concurrently considering other social factors to more directly explain this variation: source-directed attitude (Weinreich 1968, 27; Thomason 2001, 73; Lev-Ari and Peperkamp 2014), ideology regarding language contact (Poplack et al. 1988; Thomason 2001, 236), and an even broader globalist/nationalist ideological alignment. It was hypothesized that these would be stronger indexations and that any political indexation was second-order to them.

In terms of production, this second-order hypothesis led to the prediction that loanword variation would pattern along political identity, but that the other social factors considered would account for most or all of the variation observed along political identity. In terms of perception, this second-order hypothesis led to the prediction that political indexation would not be as strongly activated in an implicit matched-guise test (Lambert et al. 1960; Purnell et al. 1999) as in a metalinguistic questionnaire eliciting participants' explicit subjective evaluations (Preston 1999). Such an observation would reflect the ideological association of political identity with other indexicalities, chain-activated after the activation of more direct indexations upon more focused, metalinguistic (and what might be considered meta-social) contemplation.

The observations throughout this dissertation in many ways evince the hypothesized layered nature of the indexicality of loanword variation. In speech production, globalist/nationalist alignment is the strongest predictor, explaining much of the variation observable along political identity if the latter were the only predictor considered. However, the variation observable along political identity was not a mere byproduct: In Chapter 2, some variation of loanwords that might be considered more politically charged in the US political zeitgeist (those indexed with Arabs, Islam, and Latin America) was observable along political identity and identified as significant even after accounting for the stronger, broader effect of globalist/nationalist alignment. As discussed more thoroughly in the conclusion of Chapter 2, this suggests that the indexicality of loanword variation may have been subject to some degree of indexical reconstrual (Silverstein 2003; Eckert 2008), being extended



from the more primary indexation of globalist/nationalist alignment to the associated social factor of political identity while not seemingly any full indexical shift. In other words the political indexation is apparent in production, not as a mere byproduct but still seemingly second-order.

In perception, this layered indexicality is further evinced by a confirmed difference of perceptual indexations across implicit vs. explicit methods of elicitation. In an implicit matched-guise experiment, participants do not significantly rate the use of more source-like loanword pronunciations as more liberal-sounding, while they do rate it as sounding more globally oriented; however, in an explicit metalinguistic questionnaire, they do also rate such pronunciations as more liberal-sounding. Additionally, the use of more source-like pronunciations is indexed as prestigious, such as considering someone more upper-class or experienced in a public-speaking role like news reporting. Similarly, it appears to carry more linguistic security, in line with Boberg's (1999) findings: Those who, themselves, use more source-like pronunciations consider someone else who does to be more "intelligent" (what Preston [1999] identifies as a proxy rating for 'correctness').

Perception of loanword variation also appears to be mediated to some degree by factors of the listener. As mentioned above, those who use more source-like pronunciations feel more linguistically secure in using them. Furthermore, participants who are more liberal-identifying are also those who consider the use of more source-like pronunciations as more liberal-sounding. And, participants who, themselves, use more source-like pronunciations are also those who more strongly consider someone else who does to be globally oriented. As discussed in the conclusion of Chapter 3, these findings suggest that the use of more source-like pronunciations may be an 'in-group marker' of sorts (see Kerswill and Williams 2002; Baker et al. 2009).

How do these notions of 'in-group marker' and 'act of identity' compare to and reconcile with each other? For one, the term 'in-group marker' is used here to refer to perceptual indexation: Listeners hear a certain variant and consider the speaker using it to align with

them along one or multiple dimensions. The notion of a linguistic ‘act of identity’ (Le Page and Tabouret-Keller 1985) is used here to refer to production: Speakers are especially more likely to use a certain linguistic variant if they more strongly wish to resemble a certain group that they think using that variant will bring them to resemble. As discussed in Chapter 2, Hall-Lew et al.’s (2010) interpretations of political loanword variation amongst US politicians suggest that the variation of *Iraq* may be used as an act of political identity. The observations in Chapter 2 suggest otherwise: When looking within members of a political group, the strength with which they identify with that political group does not show a significant effect where, as an act of political identity, those more strongly holding to this identity might have more strongly exhibited the political pattern of loanword variation.

Instead, rather than an act of *political* identity, in and of itself, the use of more source-like pronunciations could be considered an act of *globalist* identity. However, this identity is an ideological sort of identity, measured here by participants aligning with statements regarding interest in foreign places, appreciation of diversity, and sensitivity to other cultures. This connects variationist and anthropological considerations that ideology can influence language behavior (e.g., Woolard 1992; Jaffe 1999; Bucholtz and Hall 2004), particularly with respect to language contact and borrowing (e.g., Poplack et al. 1988; Kroskrity 1998; Thomason 2001; Garrett 2004). We can also imagine that, if ideological differences can *result* in linguistic variation, linguistic variation may be *used to index* ideology. One might not always use linguistic variation as an *act of identity* (Le Page and Tabouret-Keller 1985) meant to signal one’s alignment with a certain group, but as an *act of ideology* meant to directly espouse a certain opinion or stance. This motivates future research to examine loanword variation as a stylistic resource for direct expression of ideology (as some have done: e.g., Hall-Lew et al. 2012; Hill 1995), particularly the broad variation of loanwords as a class and how it may be used to express broader ideologies like those regarding language contact and globalism considered here.

When combined with the observation that the use of more source-like pronunciations

is perceived as an in-group marker for those who align with both globalism and a liberal political identity, it appears that this is serving as a marker for a certain *brand* of liberal identity, signalling to others this shared, multifaceted identity. In line with sociolinguistic advancements, this encourages sociolinguists to take caution in how social groups and identities are analyzed and construed, delimited, or measured in such analysis (e.g., Rickford 1986; Cheshire 2002; Blake 2016)—in this case, political identity. As discussed in the conclusion of Chapter 2, it is worth reiterating that the observations further inform our understanding of political identity in the US political zeitgeist. These results, by way of language, suggest that those who share political identities are not homogeneous along all dimensions, even those dimensions hypothesized and observed in this dissertation to correlate with political identity. Those who are more globalist-aligning are those who are more likely to use more source-like pronunciations of loanwords in American English, even those who might not as strongly identify as politically liberal more broadly. And the fact that these social factors appear to show different effects confirms that political identity and globalist/nationalist alignment are not one and the same.

However, on the other hand, the connection of these two things should not be entirely thrown out upon recognizing within-group variation. The fact that they are not one and the same does not mean that they are completely orthogonal to and independent of each other. Given the increasing prominence and factioning of political identity in the US political zeitgeist (Abramowitz and Saunders 1998; Abramowitz 2013), its increasing association with the globalist/nationalist ideological spectrum (Jost et al. 2008), and the salience of this association in public discourse (Merry 2016), it makes sense why loanword variation would gain a political indexation. Attention should be paid to whether this indexation becomes even stronger (and/or less particular to certain word groups) or weaker, depending on the future of US sociopolitical dynamics. However, as discussed in Chapter 1 (§1.3.1), loanword variation is not a traditional sociolinguistic variable, the form of which can be analyzed as more arbitrarily sourced. Given that the surface form of a loanword is concretely rooted in

its connection to the foreign language source, the relation of these indexicalities might be predicted to remain similar as opposed to undergoing some full indexical shift. That is, one's relation to the source (source-directed attitude) or more broadly regarding languages, places, or groups considered foreign (globalist/nationalist alignment) may continue to remain a primary predictor and indexation of loanword variation, however much its political indexation may strengthen.

### **5.3 Limitations and future directions**

#### **5.3.1 Expanding to other contexts**

As mentioned in the conclusion of Chapter 3, there are many different factors that may influence how listeners perceptually index and evaluate loanword variation. And such influences are surely relevant to loanword variation in speech production as well. This dissertation therefore enlightens our understanding of loanword variation in a very narrowly constrained context: the variation of loanwords between more and less source-like pronunciations (when each is licitly available in the borrowing language's sound system) by (mostly white, Millennial-aged, college-attending) speakers of American English, and the perception of this variation by American English speakers in a Mainstream US English matrix language setting framed as a news report. While some potentially influential factors were rigorously measured and compared (e.g., identities and ideological alignments) and/or teased apart (e.g., prior exposure, [posited] source language, implicit vs. explicit elicitation), the consideration of many other factors motivates further analysis.

Loanword variation in American English was considered a prime case in which to examine how loanwords from diverse sources may pattern together as a broad class. This allows for the examination of influential factors which may be more broad and less particular to the contact language or context, like language contact ideology and globalist/nationalist alignment. However, these factors may not have the same kind of influence everywhere. There may be influences particular to the borrowing language or community. For example,

Boberg (2014) suggests that Canadian English speakers' preference for the [æ] variant of 'foreign (a)' is a reflection of Canadian anti-Americanism (with Canadians indexing [ɑ] usage as American), which may be more so the case than a reflection of other identities and/or ideologies.

Further research should also continue to question source-particular influences. For example, new or different patterns may emerge when turning more specific focus to the variation of Spanish loanwords in American English, considering that this is a more intense contact situation and additional or separate attitudes and ideologies may be involved. This would also require attention to the way such variation is used and perceived, such as taxonomizing what may constitute loanword variation like that examined here vs. 'Mock Spanish' used in a pejorative or patronizing way (Hill 1995). However, as this dissertation's findings motivate, the variation of loanwords associated with a particular source could still be just one manifestation of an even broader pattern observable across loanwords of various sources if analysis were zoomed out.

As also discussed in the conclusion of Chapter 3, language (variety) and race/ethnicity may influence both loanword variation and its perception. While the use of more source-like pronunciations of loanwords may gain a speaker capital in the global linguistic market, this English+Other capital gaining could be limited only to speakers who are white, middle-class, and/or native speakers of a variety of English judged as standard-sounding (Flores 2015). Further research could be elucidating which examines how the observations in this dissertation may or may not line up amongst speakers who are (in production) or who are presumed to be (in perception) a Person of Color and/or a speaker of a language or language variety other than that which might be considered Mainstream US English. For example, Chapter 3 examined how native monolingual speakers of American English perceive loanword variation in the speech of a news reporter speaking Mainstream US English and likely perceptually indexed as white. How might this differ if the speaker were perceived to be non-native accented or a Person of Color? Or if, rather than a news report, it were an

advertisement for a travel agency?

Even when just considering perceptual indexation, public discourse and media suggests that such specificities of the context can be influential, as well as complicated. As Lee (2009) reports about President Obama's pronunciation of *Pakistan* with the more source-like [ɑ] rather than [æ], "Pakistanis have told the White House they appreciate it." However, this may not so consistently be the case when recalling the Saturday Night Live sketch ("Antonio Mendoza" 1990), discussed in Chapter 3, in which white coworkers of a Latino man were considered ridiculous for their attempts at using a Spanish-like pronunciation of different words. This difference could be attributed to formality of the setting (reporting vs. conversation), the power/status relations of the respective interlocutors (president vs. coworkers), the semantic associations of the words themselves (Latin American placename vs. US placename: e.g., *Nicaragua* vs. *Nevada*), or many other factors. Further research is therefore motivated to examine such effects, such as different matched-guise perception experiments manipulating certain aspects while controlling others, or speech elicitation experiments doing the same.

A further extension would be analyzing how the influence of social factors identified in this dissertation extends beyond loanword variation exclusively between forms that are still each licit within the borrowing language. As also mentioned in the conclusion of Chapter 3, some of the jarring "over-pronunciations" by Antonio Mendoza's coworkers in the Saturday Night Live sketch extended beyond what might be considered the confines of the English sound system, incorporating sounds like the rhotic trill [r] (e.g., *burrito*), shifting to the even more source-like central [a] than either the [æ] or [ɑ] option (e.g., *Managua*), or leniting /d/ in a more Spanish fashion to [ð] rather than [r] as is common in American English (e.g., *enchiladas*). Part of this variation may therefore involve speakers shifting their grammars away from the confines of the borrowing language's sound system. This may pattern similarly to the variation between more and less source-like variants (each considered English-licit) observed in this dissertation. Though, it may be perceived and evaluated differently. It

may be acceptable to vary between more and less source-like pronunciations while remaining *within* the confines of the borrowing language's sound system. But, incorporating sound categories, phonetic realizations, or phonotactic structures considered otherwise illicit in the sound system could carry an entirely different meaning, it could be more polarized in its indexation, and/or it could be more sensitive to factors of the context or interlocutors. And, in terms of usage rather than perception, the patterning of this kind of variation would presumably be much more confined by direct familiarity with the source language (as observed by Poplack et al. [1988] in the variation of English-sourced loans in the speech of French-English bilinguals).

Considering this extension beyond the confines of the borrowing language's sound system also motivates considering how social factors could mediate second-language sound acquisition and explain some of the variation observed across learners. If certain speakers are identified to be better imitators of the sound forms of new words they encounter, they may also be better at imitating the forms of new sound categories or sound structures they encounter. Furthermore, attitudes and ideologies are not necessarily fixed and permanent attributes of a learner. Methods aimed at addressing attitudes and ideologies, and possibly adjusting them, could therefore have an impact on second-language sound pedagogy.

### **5.3.2 Processing considerations**

A question that arose when interpreting the results of Chapter 4 was how the social mediation of loanword adaptation/imitation interacts with processing. For one, there was an observable asymmetry between established and nonce loans regarding the effect of source-directed attitude: The variation of established loanwords patterned more strongly with explicit source-directed attitude, but the imitation of nonce loanwords was more strongly mediated by implicit bias regarding the source than by explicitly expressed attitude.

It was interpreted that this asymmetry may be due to how attitudes/biases influence shorter- vs. longer-term imitation: Implicit biases may mediate short-term processing and

imitation of nonce loanwords; but, as a word becomes more established, the speaker's pronunciation is more likely to align with their explicit source-directed attitudes. Babel's (2010) observations motivate this consideration, observing that implicit biases regarding a group are a better predictor (as compared to more explicitly manipulated attitude) of one's phonetic imitation of a speaker from that group during an online imitation task. However, the imitation task performed by Babel's participants was much more immediate, with participants hearing an utterance of a word and repeating the word right away.

The task in the current study involved hearing the word while processing a narrative in which it was uttered, with the participant's own subsequent production examined for imitation coming only after that narrative had completed and while the participant was reading a narrative aloud. This task was therefore longer-term as well as more cognitively demanding than that performed by Babel's participants, while still showing implicit bias to mediate imitation. However, this intermediary length could explain why only the gliding variable between [jə] and [iə] forms exhibited more imitation than the vowel reduction variable between [ə] and [ɛ] forms. The gliding variable might be considered less subtle and more distinct while still not contrastive (Jagers 2018), which could lead to this variable having a longer window of time during which imitation effects can be observable.

Another processing factor to consider when examining socially mediated imitation is whether the mediation takes place at the stage of perception and storage, or later at the stage of production when the word is being repeated. When imitation is observed, it can be inferred that the speaker both perceived and stored the phonetic detail and was also influenced by that phonetic detail in their own subsequent utterance. On the other hand, when observing a lack of imitation (like in the case of the vowel reduction variable, or with participants who were more nationalist-aligning), it can be unclear which stage(s) might have been less influenced by the phonetic detail. It could be that the participant both perceived and stored the phonetic detail but was less strongly influenced by that detail when subsequently reproducing the word. Or, it could be that the participant was not as attuned



to that phonetic detail in the first place or that they were less inclined to store it long-term.

Previous work on ‘priming’ has examined sociolinguistic effects on perception and storage alone. In a ‘lexical decision’ task where participants hear real vs. nonce words and are instructed to identify real words, participants are faster at recognizing a word if they heard a word with a similar sound form earlier: e.g., they are faster at recognizing *cat* as a word if they have heard *rat* earlier, as opposed to having heard *dog* beforehand (Radeau et al. 1995). This also appears to be sensitive to sociophonetic detail. Sumner and Samuel (2009) examine word processing by New Yorkers who are non-rhotic and would therefore pronounce *baker* as [ˈbekə] rather than the rhotic form [ˈbekɹ], while assuming that they are also familiar with processing the more standard and regularly encountered rhotic pronunciation. Sumner and Samuel observe that each primes the other: that previously having heard a non-rhotic pronunciation of a phonologically similar word still leads to faster recognition of the test word even when pronounced in a rhotic form, and vice versa. But, this priming effect is stronger when matched for sociophonetic detail: Participants recognize the test word even faster when the prime word matches the test word along this variable of rhoticity, even though the test and exposure utterances are by different speakers of different sexes (and therefore not the exact same audio signal) and even when the time window between exposure and test is a longer 20-30min period.

Sumner and Samuel’s observations suggest that phonetic detail can be stored and influence future processing over a rather long period of time. But, this could still be socially mediated by one’s attitudes, as the current study (and Babel 2010) observes the full process of imitation to be. To isolate and test whether the socially mediated imitation observed in this study was mediated at the stage of perception and storage, future work could use a paradigm similar to that used in Chapter 4; but, after listening to the exposure story, the participant could simply be tested on whether they remember the pronunciation they heard used by the speaker, rather than eliciting a subsequent utterance and examining how closely it resembles the pronunciation they heard used. If a similar effect of social factors is present

where more nationalist-aligning participants exhibit weaker influence of the exposure form, this could suggest that their weaker imitation during production was rooted in a weaker attunement to or long-term storage of that detail. If, on the other hand, they still appear to remember the phonetic detail of the utterance they heard, this might suggest that it is more so their subsequent production and replication of that detail that is socially mediated.

### **5.3.3 Loanword status as an influence**

As argued in Chapter 1 (§1.2.3), loanword status may play an influential role in language processing and variation. As each study within this dissertation has evinced, it should definitely continue to be considered in future research regarding language contact and sociolinguistic variation. The variation of loanwords from multiple sources can pattern together as a broad class, most strongly predicted by and indexed with a similarly broad social factor of global alignment and orientation. But, when comparing the imitation of nonce words framed as loans with the imitation of nonce words framed as non-loans in Chapter 4, they were observed to pattern in parallel. It was hypothesized that speakers who were more globalist-aligning might show stronger imitation of nonce loans relative to their imitation of nonce non-loans, in comparison with those who were more nationalist-aligning. Instead, more globalist-aligning speakers showed stronger imitation of nonce words across the board, including those framed as non-loans. Therefore, loanwords may pattern together as a broad class in variation, but loanword status in and of itself might not mediate processing and imitation in a special way. However, this idea should not yet be entirely thrown out. Future work should continue to test the effect of loanword status on the processing of new words, given that much phonological analysis suggests loanword status to play a role in determining the sound forms of at least more established words (e.g., Saciuk 1969; Itô and Mester 1999; Smith 2006; 2009b).

The study of sound form imitation in Chapter 4 was designed with the express intent of teasing apart loanword status from phonological non-nativeness, examining whether loanword status in and of itself might mediate the degree of imitation. These findings may

therefore suggest that loanword status, and even its interaction with social factors, does not influence the general degree of imitation during initial processing. However, there may be particular sounds or phonotactics that are especially favored when adapting a loanword and which may be detectable by a similar loan/non-loan experimental paradigm.

One phenomenon suggests that the effect of loanword status on sound processing may induce certain phonotactic preferences. When adapting a loanword to a form that is more pronounceable by speakers of the borrowing language, sound sequences in the source form that are otherwise disallowed in the borrowing language show a preference of being resolved by the insertion of new sounds, rather than other options like deletion or significant modification of the source sounds: Paradis and LaCharité's (1997) Preservation Principle. There are competing theories regarding this observed preference, suggesting that either in perception (Peperkamp et al. 2008) or production (Davidson 2007) the form is forced to comply with the borrowing language's sound system. These theories tend to rely on the non-nativeness of the sound sequence being borrowed, such as Davidson's (2007) observation that native English speakers tend to reproduce consonant clusters that would otherwise be illicit in English ( $*/CC/$ ) with an excrescent schwa between the two consonants ( $[C^{\text{ə}}C]$ ) as a result of trying to produce both but mis-timing their coordination, then resulting in a full phonologization of that excrescent schwa ( $[C^{\text{ə}}C] \rightarrow /C^{\text{ə}}C/$ ) further down the path of a loanword's dissemination.

However, it may be the case that loanword status itself induces this by way of Weinberger's (1994) Recoverability Principle regarding second-language sound acquisition: With a word being foreign, one may be unsure about what sounds or sound features are expendable in the source language; because of this, they will preserve as much as possible to avoid losing details potentially crucial to recognizing the word, even if this involves inserting some new material to make it more pronounceable. Like in the experimentation conducted in Chapter 4, loanword status can be teased apart from phonological non-nativeness to examine whether it induces such a preference. Loanword status may increase listeners' inclination

to transcribe or repeat an exposure pronunciation ambiguous between two options with the option that has more material (e.g., interpreting [C<sup>o</sup>C] as /CəC/ rather than /CC/), even if the option with less material is still compliant with the borrowing language's sound system (e.g., nonce word /snum/~/sənúm/ in English). And, we might imagine that this could also be socially mediated, with speakers who are more nationalist-aligning being less concerned about 'recoverability' and therefore exhibiting less of an effect of loanword status in this linguistic behavior.

Another special behavior of loanwords suggests that there are certain sound forms that get indexed as 'foreign-sounding' and preferred in loanword adaptation, such as Americans more often positing that the back [ɑ] vowel is the source pronunciation of a low vowel than the front [æ] (Boberg 1999). These indexations sometimes appear to override replication of the source form, such as the use of [ɜ] in words like *Beijing* and *Taj Mahal* even when [ɔ̃] would be a closer replication of those source forms and is a readily available sound in the English inventory. Such sound preferences should be rigorously experimentally tested. Sounds may be preferred because they are available while infrequent in the borrowing language's sound inventory (e.g., [ɜ]), or because of spelling-to-sound mappings that are less common among older native words (e.g., <a> → [ɑ]). The salience of the sound difference may also play a role, such as *Bach* being more readily variable between [bak] and the more German-like [bax] pronunciations, while *Mexico* pronounced with [x] rather than the standardly Anglicized [ks] seems less frequent and more marked (in spite of /x/ being considered only a peripheral member of the English sound inventory, if even that). Experimentation like that described above (e.g., nonce words, loan vs. non-loan framings, testing for preference in processing ambiguous exposure forms) can test and tease apart some of these influences.

There also is an open question of how such conventions are formed. It seems likely that they may be socially rooted. For example, Janda et al. (1994) suggest that cases like *Bei[ɜ]ing* and *Ta[ɜ] Mahal* are an overgeneralization of the French <j> → [ɜ] spelling-to-sound mapping, with this being applied more generally to loanwords beyond those that are

French-sourced. But which contact languages' patterns get overgeneralized and why? This may be addressed by examining the borrowing language's relative intensity of contact with different source languages. For example, Kim (2017) observes that older Korean loanwords borrowed from French show word-final vowel insertion more than English-sourced loanwords. But, Kim observes that this is undergoing a change, where newer loanwords both French- and English-sourced more frequently are adapted the way English-sourced loanwords are: with less frequent vowel insertion. Kim suggests that the chronological increase in direct Korean-English contact has led to this change in how loanwords as a broad class are adapted, with the adaptation used for English-sourced loanwords gaining preference. Generalizations like this could also be influenced by other social factors less directly related to contact intensity, such as the degree of prestige ascribed to a particular contact language relative to others.

Furthermore, the study of loanword adaptation has provided insight to the nature of phonological language change. Studies regarding loanword phonology (e.g., Kiparsky 1968; Saciuk 1969; Itô and Mester 1999) suggest that epochs of extensive borrowing in one language of words from a particular language or group of languages can significantly change the phonology of the borrowing language. Faithfulness to the sounds or phonotactics of many source forms of loanwords containing them can become regularized, leading to a lexical stratum of phonological exceptionality. Then, in later epochs of language contact and new corresponding strata of lexical borrowings, what might previously have been considered phonological exceptionality specific to a particular lexical stratum applies to subsequent borrowings as well, suggesting a broader change to the grammar. The current study's sociolinguistic considerations may illuminate how such exceptional faithfulness may become regularized. As mentioned in Chapter 1, the historic change of English phonology to a phonemic, contrastive status between /f/ and /v/ (rather than the previously allophonic status), which was attributable to French loanwords borrowed following the Norman Conquest (Smith 2009a), may have been especially driven by the high degree of prestige ascribed to French (Baugh and Cable 2013). The current study's findings have provided us with a better sense of *who*

might lead to such change. Speakers who exhibit more faithfulness to the sound form of a new loanword may also be more likely to develop strata of phonological exceptionality and may therefore be the hypothesized leaders of broader contact-induced language change.

# Appendices

## A Likert questionnaire

The following statements were provided, along with a 7-point Likert response scale ranging from “strongly agree” to “strongly disagree”. These were presented one per screen, with the questionnaire auto-advancing to the next upon a participant’s selection. Order was semi-randomized with attention to spacing out statements regarding the same factor. Order number is provided, while statements below are grouped the factor of interest.

FACTOR	DIRECTION	ORDER	STATEMENT
political identity	liberal	3	My political views are usually liberal.
		16	I usually agree with the Democrat party.
		32	I have a lot of Democrat friends.
	conservative	28	My political views are usually conservative.
		9	I usually agree with the Republican party.
		20	I have a lot of Republican friends.
(not included in analysis)		24	I often discuss my opinions about politics.

globalist/nationalist alignment	globalist	14	I consider myself sensitive to other cultures.
		25	I would avoid identifying myself as an American if I were traveling in a foreign country.
		1	I often fantasize traveling to new places around the world.
		36	Immigrants who move to the US shouldn't be expected to adopt American cultural practices.
	nationalist	13	I don't keep myself up to date much regarding international current events.
		29	From what I hear in the news about other countries, I feel fortunate to be living in the US.
		19	I'm not much interested in traveling, national or abroad.
		4	The negative stereotypes that get attached to Americans are often unfounded or wrong.
language contact ideology	more receptive	18	It's important to learn some of the language of a place one travels to.
		2	The US should have stronger foreign language education in schools.
		35	I enjoy learning other languages.
	less receptive	7	Immigrants who move to the US should be required to learn English.
		30	English should be the official language of the US.
		23	It's annoying to me when someone uses a word from another language while speaking English.



place-directed attitude	Iraq-favoring	33	I would like to travel to Iraq if I had the chance.
		15	If invited, I would be willing to go to an Iraqi festival.
	Iraq-disfavoring	8	I would consider Iraq a dangerous place to travel to.
		26	I am not particularly interested in learning about Iraqi culture.
	Indonesia-favoring	12	I would like to travel to Indonesia if I had the chance.
		6	If invited, I would be willing to go to an Indonesian festival.
	Indonesia-disfavoring	21	I would consider Indonesia a dangerous place to travel to.
		37	I am not particularly interested in learning about Indonesian culture.
prescriptivism	prescriptivist	5	It's annoying when someone doesn't use proper English grammar.
		11	Customer service representatives shouldn't have strong foreign accents.
	descriptivist	34	I enjoy hearing different regional accents of English.
		17	People with strong foreign accents are just as likely to be smart as people without accents.
mobility	more mobile	22	I have traveled to a diverse range of places, compared to most.
		31	I have traveled outside of the US with my family many times.
	less mobile	27	I haven't had the money that would allow me to travel much.
		10	I haven't had the opportunity to travel much.

## B Step-up modeling, Study 2

{*Iran, Iraq*} variation:

Base:

```
glmer(variant ~ (1|word),
      data = data[data$political.identity <= 4,], family = "binomial")
```

Results:

Fixed effect	Estimate	Std. Err.	z value
(intercept)	0	.163	0

Step 1 (lone-predictor models):

```
glmer(variant ~ fixed.effect + (1|word),
      data = data[data$political.identity <= 4,], family = "binomial")
```

Results:

Fixed effect	Estimate	Std. Err.	z value	Pr(> z )
globalist/nationalist alignment	-.921	.216	-4.26	2.04e-5
language contact ideology	-.551	.187	-2.956	.0031
Middle East-directed attitude (Likert)	-.346	.17	-2.036	.042
political identity	-.215	.166	-1.296	.195
Middle East-directed attitude (IAT)	-.181	.165	-1.095	.273
mobility	.091	.098	.936	.349

Step 2:

```
glmer(variant ~ fixed.effect
      + globalist.nationalist.alignment
      + (1|word),
      data = data[data$political.identity <= 4,], family = "binomial")
```

Results:

Fixed effect	Estimate	Std. Err.	z value	Pr(> z )	Pr(>Chisq)
political identity	.467	.228	2.042	.041	.036
Middle East-directed attitude (Likert)	-.275	.181	-1.523	.128	.123
Middle East-directed attitude (IAT)	-.272	.185	-1.472	.141	.136
language contact ideology	.194	.287	.675	.499	.498
mobility	.00014	.107	.001	.999	.999

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

### Step 3:

```
glmer(variant ~ fixed.effect
      + globalist.nationalist.alignment + political.identity
      + (1|word),
      data = data[data$political.identity <= 4,], family = "binomial")
```

Results:

Fixed effect	Estimate	Std. Err.	z value	Pr(> z )	Pr(>Chisq)
Middle East-directed attitude (Likert)	-.329	.186	-1.77	.077	.073
Middle East-directed attitude (IAT)	-.235	.189	-1.243	.214	.21
mobility	-.062	.116	-.532	.595	.593
language contact ideology	-.022	.311	-.072	.943	.9428

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

### Final model:

```
glmer(variant ~ globalist.nationalist.alignment + political.identity
      + (1|word),
      data = data[data$political.identity <= 4,], family = "binomial")
```

AIC	BIC	logLik	deviance	df.resid		
188	200.1	-90	180	146		
Scaled residuals:						
Min	1Q	Median	3Q	Max		
-1.731	-0.801	0.043	0.908	1.489		
Fixed effects:		Estimate	Std. Err.	z value	Pr(> z )	Pr(>Chisq)
(intercept)		-.004	.179	-.021	.983	
globalist/nationalist alignment		-1.236	.273	-4.506	6.62e-6	3.081e-7
political identity		.467	.228	2.042	.041	.0361

\*Pr(>Chisq) comes from Chi-square ANOVA test between full model and model excluding component of interest.

### Full set of target loans:

#### Base:

```
glmer(variant ~ (1|word),  
      data = data[data$political.identity <= 4,], family = "binomial")
```

Results:

Fixed effect	Estimate	Std. Err.	z value
(intercept)	-.253	.46	-.551

#### Step 1 (lone-predictor models):

```
glmer(variant ~ fixed.effect  
      + (1|word),  
      data = data[data$political.identity <= 4,], family = "binomial")
```

Results:

Fixed effect	Estimate	Std. Err.	z value	Pr(> z )
globalist/nationalist	-.475	.099	-4.758	1.96e-6
alignment				
language contact	-.296	.095	-3.11	.0019
ideology				
Middle East-directed	-.212	.094	-2.257	.024
attitude (Likert)				
political identity	-.129	.093	-1.399	.162
mobility	.057	.055	1.036	.3
Middle East-directed	.014	.093	.093	.88
attitude (IAT)				

#### Step 2:

```
glmer(variant ~ fixed.effect  
      + globalist.nationalist.alignment  
      + (1|word),  
      data = data[data$political.identity <= 4,], family = "binomial")
```

Results:

Fixed effect	Estimate	Std. Err.	z value	Pr(> z )	Pr(>Chisq)
political identity	.253	.121	2.092	.036	.036
Middle East-directed attitude (Likert)	-.149	.096	-1.559	.119	.119
language contact ideology	.142	.149	0.951	.341	.342
Middle East-directed attitude (IAT)	-.032	.096	-0.331	.741	.741
mobility	.005	.057	0.081	.935	.936

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

### Step 3:

```
glmer(variant ~ fixed.effect
      + globalist.nationalist.alignment + political.identity
      + (1|word),
      data = data[data$political.identity <= 4,], family = "binomial")
```

Results:

Fixed effect	Estimate	Std. Err.	z value	Pr(> z )	Pr(>Chisq)
Middle East-directed attitude (Likert)	-.174	.096	-1.806	.071	.071
mobility	-.025	.059	-0.415	.678	.679
language contact ideology	.037	.159	.231	.817	.818
Middle East-directed attitude (IAT)	-.007	.097	-0.072	.942	.943

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

### Final model:

```
glmer(variant ~ globalist.nationalist.alignment + political.identity
      + (1|word),
      data = data[data$political.identity <= 4,], family = "binomial")
```

AIC	BIC	logLik	deviance	df.resid		
683	700.6	-337.5	675	598		
Scaled residuals:						
Min	1Q	Median	3Q	Max		
-2.271	-0.795	-0.234	-.708	4.304		
Fixed effects:		Estimate	Std. Err.	z value	Pr(> z )	Pr(>Chisq)
(intercept)		-.263	.481	-0.546	0.586	
globalist/nationalist alignment		-.636	.128	-4.997	5.83e-7	2.268e-7
political identity		.253	.121	2.092	.036	.036

\*Pr(>Chisq) comes from Chi-square ANOVA test between full model and model excluding component of interest.

## C Step-up modeling, Study 3

{*Iran, Iraq*} variation:

Base:

```
lmer(num.of.divergences ~ (1|word),
      data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value
(intercept)	1.526	.057	26.61

Step 1 (lone-predictor models):

```
lmer(num.of.divergences ~ fixed.effect + (1|word),
      data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
globalist/nationalist alignment	.139	.057	2.472	.014
political identity	.125	.057	2.205	.028
mobility	.097	.057	1.701	.089
prescriptivism	.073	.057	1.287	.197
language contact ideology	.065	.057	1.142	.252
Middle East-directed attitude (IAT)	-.034	.057	-0.587	.555
Middle East-directed attitude (Likert)	.007	.057	0.121	.903

\*Pr(>Chisq) comes from Chi-square ANOVA test between model from previous step and an identical model additionally including the fixed effect of interest in the current step.

Step 2:

```
lmer(num.of.divergences ~ fixed.effect
      + globalist.nationalist.alignment
      + (1|word),
      data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
prescriptivism	.129	.059	2.207	.027
political identity	.089	.059	1.508	.129
mobility	.082	.057	1.44	.148
Middle East-directed attitude (IAT)	-.024	.057	-0.427	.667
Middle East-directed attitude (Likert)	-.009	.057	-0.169	.865
language contact ideology	-.01	.066	-0.151	.879

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

### Step 3:

```
lmer(num.of.divergences ~ fixed.effect
      + globalist.nationalist.alignment + prescriptivism
      + (1|word),
      data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
political identity	.107	.059	1.802	.069
mobility	.083	.056	1.477	.137
language contact ideology	.054	.071	0.766	.439
Middle East-directed attitude (IAT)	-.021	.056	-0.381	.7
Middle East-directed attitude (Likert)	.014	.057	0.239	.809

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

### Final model:

```
lmer(num.of.divergences ~ globalist.nationalist.alignment + prescriptivism
      + (1|word),
      data = data)
```

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.365	-0.56	0.464	0.678	1.775

Fixed effects:	Estimate	Std. Err.	t value	Pr(>Chisq)
(intercept)	1.525	.056	27.218	
globalist/nationalist alignment	.181	.059	3.062	.0023
prescriptivism	.129	.059	2.207	.027

\*Pr(>Chisq) comes from Chi-square ANOVA test between full model and model excluding component of interest.

### Full set of target loans:

#### Base:

```
lmer(num.of.divergences ~ (1|word),
      data = data)
```

#### Results:

Fixed effect	Estimate	Std. Err.	t value
(intercept)	1.204	.163	7.39

#### Step 1 (lone-predictor models):

```
lmer(num.of.divergences ~ fixed.effect + (1|word),
      data = data)
```

#### Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
globalist/nationalist alignment	.069	.014	4.907	1.02e-6
mobility	.037	.014	2.621	.009
language contact ideology	.036	.014	2.565	.01
political identity**	.054	.019	2.771	.01
prescriptivism	-.005	.014	-0.326	.744
Middle East-directed attitude (Likert)	-.004	.014	-0.259	.796
Middle East-directed attitude (IAT)	-.0006	.014	-0.046	.964

\*Pr(>Chisq) comes from Chi-square ANOVA test between model from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a "+ (fixed.effect|word)" term was compared to a model without, to test if including a random slope per word was motivated. For this fixed effect, the inclusion of a random slope was found to significantly improve model fit.

#### Step 2:

```
lmer(num.of.divergences ~ fixed.effect
      + globalist.nationalist.alignment
      + (1|word),
```



```
data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
mobility	.029	.014	2.067	.039
political identity**	.036	.02	1.799	.077
prescriptivism	.019	.015	1.301	.193
Middle East-directed attitude (Likert)	-.011	.014	-0.812	.661
Middle East-directed attitude (IAT)	.004	.014	0.283	.777
language contact ideology	.001	.016	0.048	.962

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ (fixed.effect|word)” term was compared to a model without, to test if including a random slope per word was motivated. For this fixed effect, the inclusion of a random slope was found to significantly improve model fit.

Step 3:

```
lmer(num.of.divergences ~ fixed.effect
      + globalist.nationalist.alignment + mobility
      + (1|word),
      data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
political identity**	.044	.02	2.168	.036
prescriptivism	.019	.015	1.33	.183
Middle East-directed attitude (Likert)	-.004	.015	-0.298	.765
language contact ideology	.002	.016	0.148	.883
Middle East-directed attitude (IAT)	.002	.014	0.144	.885

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ (fixed.effect|word)” term was compared to a model without, to test if including a random slope per word was motivated. For this fixed effect, the inclusion of a random slope was found to significantly improve model fit.

Step 4:

```
lmer(num.of.divergences ~ fixed.effect
      + globalist.nationalist.alignment + mobility + political.identity
      + (1+political.identity|word),
      data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
prescriptivism	.025	.015	1.697	.089
language contact ideology	-.012	.017	-0.696	.486
Middle East-directed attitude (Likert)	-.004	.015	-0.259	.795
Middle East-directed attitude (IAT)	.001	.014	0.074	.941

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ (fixed.effect|word)” term was compared to a model without, to test if including a random slope per word was motivated. For this fixed effect, the inclusion of a random slope was found to significantly improve model fit.

### Final model:

```
lmer(num.of.divergences ~
  + globalist.nationalist.alignment + mobility + political.identity
  + (1+political.identity|word),
  data = data)
```

### Results:

Scaled residuals:

Min	1Q	Median	3Q	Max
-5.321	-0.532	0.269	0.634	1.909

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
(intercept)	1.204	.163	7.39	
globalist/nationalist alignment	.049	.015	3.356	.0008
mobility	.038	.014	2.65	.008
political identity**	.044	.02	2.168	.036

\*Pr(>Chisq) comes from Chi-square ANOVA test between full model and model excluding component of interest.

\*\*For each fixed effect, a model including a “+ (fixed.effect|word)” term was compared to a model without, to test if including a random slope per word was motivated. For this fixed effect, the inclusion of a random slope was found to significantly improve model fit.

```
coef(final.model)$word[order(-coef(final.model)$word$political.identity),
  c("(Intercept)","political.identity")]
```

Fixed effect	(Intercept)	political.identity
Muslim	2.756	0.141
Iraq	1.558	0.098
Chile	1.195	0.094
tamales	1.688	0.074
Paraguay	1.715	0.07
Iran	1.462	0.069
foyer	1.77	0.055
Hawaii	0.969	0.024
bruschetta	0.928	0.02
spiel	0.667	0.018
tsunami	0.968	0.011
Nevada	0.719	0.009
Quebec	0.807	0.006
genre	0.337	-0.005
Tokyo	0.524	-0.018

## D Matched-guise news reports

### Target news report stimulus containing variable placenames of interest:

The US Department of Transportation has approved a foreign air carrier permit for a subsidiary of Iceland Airlines to fly to the United States. The airline has recently formed multiple subsidiaries such as those based in Tokyo, Japan and Budapest, Hungary. The particular subsidiary of interest here is one based in Shanghai, in China.

Late last year a tentative permit was approved while further negotiation took place. Some criticism was raised by opponents, considering the subsidiary a way for the airline to circumvent the stricter labor laws of Iceland. Opponents cited the existence of further international subsidiaries, and also the practice of hiring crews under contracts based in Paraguay and, more recently, further South American countries like Chile and Colombia. However, US officials said that such concerns provided no legal basis for rejecting such an application.

Iceland Airlines has been expanding significantly, recently offering flights to a broad range of new destination countries such as Tanzania, New Zealand, Iraq, and Pakistan. This also includes new landing sites within already offered destination countries, such as flights to Canada now also landing in Quebec, and flights to Russia landing in Saint Petersburg. This new permit granted by the US Department of Transportation will further strengthen the airline's competitive standing.

\*Based on pre-existing news report (Baker 2016). Some details and proper names purposefully changed.

### Filler (preceding) news report stimulus:

South Korea recently announced revised requirements for a new generation of spaceflight vehicles it plans to purchase. This past Friday, South Korea's Aerospace Research Agency issued a statement after a meeting of its top administrators. This statement included plans to buy 30 high-capability launch vehicles during the time span of 2018 to 2024. This is a significant advancement for South Korean aerospace research, further reaching a caliber like that of the US, Russia, France, and the UK.

Original contenders for the purchase included bids from American, British, and German manufacturing companies. However, the heightened requirements will effectively leave American manufacturer Lockheed Martin's Atlas 5 N22 as the only viable bidder. The announcement could bolster the N22's growing dominance in the global market, which could be seen as an intended outcome given the tight relationship between South Korea and the US.

However, the N22 is also known to have been purchased recently by many other countries such as Australia, Israel, Japan, and the Netherlands. There are further reports that South Korea may seek smaller bids for additional spaceflight vehicles and equipment of less stringent capability requirements. These are expected to garner competing bids from multiple non-American manufacturing companies.

\*Based on pre-existing news report (Choe 2013). Some details and proper names purposefully changed.

## E Modeling effects of individual factors on sociolinguistic perception

### Matched-guise:

```
lm(dependent.variable ~ exposure.condition
  + Iraq.self.rept : exposure.condition
  + polit.ident.self.rept : exposure.condition,
  data = data)
```

### Results:

“MSL” = more source-like, “LSL” = less source-like,  
 “intercept” = expos.cond.LSL & Iraq.self.rept.rack)

Dependent variable:		pleasantness			
Coefficient	Estimate	Std. Err.	t value	Pr(> t )	
(intercept)	3.174	.259	12.209		
expos.cond.MSL	-.149	.374	-.399	.69	
expos.cond.LSL : Iraq.self.rept.rock	.024	.342	.07	.944	
expos.cond.MSL : Iraq.self.rept.rock	-.109	.285	-.381	.704	
expos.cond.LSL : polit.ident.self	.073	.068	1.088	.277	
expos.cond.MSL : polit.ident.self	.07	.067	1.046	.296	

Dependent variable:		experience			
Coefficient	Estimate	Std. Err.	t value	Pr(> t )	
(intercept)	4.228	.224	18.857		
expos.cond.MSL	.39	.323	1.21	.227	
expos.cond.LSL : Iraq.self.rept.rock	-.557	.295	-1.887	.059	
expos.cond.MSL : Iraq.self.rept.rock	.117	.246	.475	.635	
expos.cond.LSL : polit.ident.self	.141	.058	2.422	.016	
expos.cond.MSL : polit.ident.self	.083	.058	1.43	.153	

Dependent variable:		intelligence			
Coefficient	Estimate	Std. Err.	t value	Pr(> t )	
(intercept)	3.227	.249	12.987		
expos.cond.MSL	.006	.358	.018	.986	
expos.cond.LSL : Iraq.self.rept.rock	.202	.327	.618	.537	
expos.cond.MSL : Iraq.self.rept.rock	-.241	.273	-.883	.378	
expos.cond.LSL : polit.ident.self	.049	.065	.758	.449	
expos.cond.MSL : polit.ident.self	.018	.064	.277	.782	

Dependent variable:		knowl.re.world.affairs				
Coefficient		Estimate	Std. Err.	t value	Pr(> t )	
(intercept)		3.328	.241	13.834		
expos.cond.MSL		-.053	.346	-.154	.877	
expos.cond.LSL : Iraq.self.rept.rock		.025	.316	.08	.936	
expos.cond.MSL : Iraq.self.rept.rock		-.092	.264	-.348	.728	
expos.cond.LSL : polit.ident.self		.017	.063	.269	.788	
expos.cond.MSL : polit.ident.self		-.007	.062	-.113	.91	

Dependent variable:		multilingualism				
Coefficient		Estimate	Std. Err.	t value	Pr(> t )	
(intercept)		3.515	.235	14.948		
expos.cond.MSL		.397	.338	1.172	.242	
expos.cond.LSL : Iraq.self.rept.rock		-.227	.309	-.732	.464	
expos.cond.MSL : Iraq.self.rept.rock		.061	.258	.236	.814	
expos.cond.LSL : polit.ident.self		-.043	.061	-.709	.479	
expos.cond.MSL : polit.ident.self		.048	.061	.792	.429	

Dependent variable:		local.global.station				
Coefficient		Estimate	Std. Err.	t value	Pr(> t )	
(intercept)		3.516	.228	15.44		
expos.cond.MSL		.478	.328	1.457	.146	
expos.cond.LSL : Iraq.self.rept.rock		-.643	.299	-2.148	.032	
expos.cond.MSL : Iraq.self.rept.rock		.122	.25	.488	.626	
expos.cond.LSL : polit.ident.self		.164	.059	2.777	.0058	
expos.cond.MSL : polit.ident.self		.086	.059	1.452	.147	

Dependent variable:		political.leaning				
Coefficient		Estimate	Std. Err.	t value	Pr(> t )	
(intercept)		3.558	.152	23.431		
expos.cond.MSL		-.149	.219	-.685	.494	
expos.cond.LSL : Iraq.self.rept.rock		-.155	.199	-.775	.439	
expos.cond.MSL : Iraq.self.rept.rock		.169	.167	1.012	.312	
expos.cond.LSL : polit.ident.self		.07	.039	1.78	.076	
expos.cond.MSL : polit.ident.self		.051	.039	1.294	.196	

Dependent variable:		comprehens				
Coefficient		Estimate	Std. Err.	t value	Pr(> t )	
(intercept)		5.222	.221	23.619		
expos.cond.MSL		.442	.318	1.387	.166	
expos.cond.LSL : Iraq.self.rept.rock		.036	.291	.125	.901	
expos.cond.MSL : Iraq.self.rept.rock		-.131	.243	-.538	.591	
expos.cond.LSL : polit.ident.self		.051	.058	.879	.38	
expos.cond.MSL : polit.ident.self		-.008	.057	-.148	.882	

Dependent variable:		reliability			
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )
(intercept)	3.596	.271		13.295	
expos.cond.MSL	-.23	.389		-.592	.554
expos.cond.LSL : Iraq.self.rept.rock	.017	.356		.047	.962
expos.cond.MSL : Iraq.self.rept.rock	-.125	.297		-.422	.673
expos.cond.LSL : polit.ident.self	-.005	.07		-.067	.947
expos.cond.MSL : polit.ident.self	.009	.07		.131	.896

### Metalinguistic ratings:

```
lm(dependent.variable ~ exposure.condition
  + Iraq.self.rept : exposure.condition
  + polit.ident.self.rept : exposure.condition,
  data = data)
```

### Results:

("MSL" = more source-like, "LSL" = less source-like,  
 "intercept" = expos.cond.LSL & Iraq.self.rept.rock)

Dependent variable:		friendliness			
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )
(intercept)	4.324	.156		27.683	
expos.cond.MSL	.339	.227		1.495	.136
expos.cond.LSL : Iraq.self.rept.rock	-.221	.176		-1.257	.209
expos.cond.MSL : Iraq.self.rept.rock	0.211	.199		-1.056	.292
expos.cond.LSL : polit.ident.self	.041	.04		1.02	.309
expos.cond.MSL : polit.ident.self	.019	.041		.465	.642

Dependent variable:		intelligence			
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )
(intercept)	3.789	.191		19.887	
expos.cond.MSL	-.664	.276		-2.403	.017
expos.cond.LSL : Iraq.self.rept.rock	.619	.215		2.88	.0042
expos.cond.MSL : Iraq.self.rept.rock	-.017	.243		-.071	.944
expos.cond.LSL : polit.ident.self	-.056	.049		-1.14	.255
expos.cond.MSL : polit.ident.self	.004	.05		.078	.938

Dependent variable:		education			
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )
(intercept)	4.033	.201		20.113	
expos.cond.MSL	-.954	.291		-3.28	.0011
expos.cond.LSL : Iraq.self.rept.rock	.377	.226		1.668	.096
expos.cond.MSL : Iraq.self.rept.rock	-.328	.256		-1.28	.201
expos.cond.LSL : polit.ident.self	-.064	.052		-1.241	.215
expos.cond.MSL : polit.ident.self	.017	.053		.328	.743

Dependent variable:		social.class				
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )	
(intercept)	3.478	.168		20.656		
expos.cond.MSL	1.071	.244		4.383	1.5e-5	
expos.cond.LSL : Iraq.self.rept.rock	-.307	.189		-1.616	.107	
expos.cond.MSL : Iraq.self.rept.rock	.151	.215		.703	.482	
expos.cond.LSL : polit.ident.self	.125	.044		2.857	.0045	
expos.cond.MSL : polit.ident.self	-.021	.045		-.469	.639	

Dependent variable:		humble.pretentious				
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )	
(intercept)	3.703	.177		20.922		
expos.cond.MSL	.289	.257		1.124	.262	
expos.cond.LSL : Iraq.self.rept.rock	.083	.199		.414	.679	
expos.cond.MSL : Iraq.self.rept.rock	-.052	.226		-.231	.818	
expos.cond.LSL : polit.ident.self	-.003	.046		-.058	.954	
expos.cond.MSL : polit.ident.self	.001	.047		.016	.988	

Dependent variable:		polit.ident				
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )	
(intercept)	4.273	.175		24.398		
expos.cond.MSL	-1.218	.254		-4.794	2.31e-6	
expos.cond.LSL : Iraq.self.rept.rock	.416	.198		2.106	.036	
expos.cond.MSL : Iraq.self.rept.rock	-.251	.224		-1.119	.264	
expos.cond.LSL : polit.ident.self	-.015	.045		-.335	.738	
expos.cond.MSL : polit.ident.self	.172	.046		3.717	.00023	

Dependent variable:		openmindedness				
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )	
(intercept)	3.832	.188		20.341		
expos.cond.MSL	.877	.273		3.208	.0015	
expos.cond.LSL : Iraq.self.rept.rock	-.471	.212		-2.218	.027	
expos.cond.MSL : Iraq.self.rept.rock	.212	.241		.883	.378	
expos.cond.LSL : polit.ident.self	.067	.049		1.37	.171	
expos.cond.MSL : polit.ident.self	-.027	.049		-.547	.584	

Dependent variable:		multilingualism				
Coefficient	Estimate	Std.	Err.	t value	Pr(> t )	
(intercept)	2.887	.247		11.703		
expos.cond.MSL	1.949	.358		5.449	8.94e-8	
expos.cond.LSL : Iraq.self.rept.rock	-.401	.278		-1.44	.151	
expos.cond.MSL : Iraq.self.rept.rock	-.398	.315		-1.265	.207	
expos.cond.LSL : polit.ident.self	.069	.064		1.089	.277	
expos.cond.MSL : polit.ident.self	-.023	.065		-.345	.73	



## **F Stories and sequels**

### **Trial 1 (“Exploring”) exposure story, US framing:**

Last year, I was the lead on the biggest mission my cave exploration team had ever taken on. We would be the first team to officially map Sloxy cave, located in Northern Utah. Allison was my assistant who traveled with me and helped me prepare.

My hometown being just an hour’s drive away, I had been to the cave once before when I was a kid. It was actually what first got me into cave exploring, with its hennia formations that are unique to this region. So I was excited to return to it to try to map it.

We met with Mr. Clote, a local resident who knows the cave well and who would be joining me. He gave me a semave to use on the mission, which is a kind of echolocation tool he uses to gauge the distance of a chasm.

The first day was just me and Mr. Clote. After going through some long, winding tunnels, the cave opened up to a pretty open cavern. We spent time taking pictures of some of the different formations we observed. There was one particularly large hennia formation in the center of the cavern. That helped serve as a nice anchoring point when we then worked on taking measurements for accurate scaling of the map.

Before heading back out at the end of the day, we made sure to note all of the different tunnels and chasms possibly leading to other areas. There was one dauntingly narrow tunnel. And when we used the semave to gauge it, it seemed like it was quite long as well. It was too narrow for either of us to fit into, but Allison might. One thing was sure: There was a lot more of Sloxy cave remaining to be explored in the weeks to come.

### **Trial 1 (“Exploring”) sequel, US framing:**

Allison was intrigued to hear about how the first day in Sloxy cave had gone. She had read about the hennia formations and wanted a chance to see them in person. These were probably her biggest reason for wanting to go on this mission.

When Mr. Clote described how long the initial tunnels were, she thought that sounded kind of boring. But she was excited to hear about the really narrow opening they had found. They had used the semave for an estimation, suggesting that the tunnel was quite long. Allison had become quite an expert at narrow passages, so she jumped at the chance when they said they’d need her to do that one solo.

Then they reviewed their measurements and photos to start creating a map. Sloxy cave definitely photographed well. The hennia pictures they had taken looked amazing. She was excited to see them for herself. And the measurements from the main cavern were surprisingly large.

At the end of the day, Allison was glad to have joined the mission. She also had lots of prep to do. She would need to use the semave for her solo mission, so she’d make sure to get Mr. Clote to teach her how. And she’d need to make sure she had the best gear for tight spaces.

### **Trial 1 (“Exploring”) exposure story, Iraq framing:**

Last year, I was the lead on the biggest mission my cave exploration team had ever taken on. We would be the first team to officially map Slaqsi cave, located in Western Iraq. Allison was my assistant who traveled with me and helped me prepare, and I translated between her and the local residents.

My hometown being just an hour’s drive away, I had been to the cave once before when I was a kid. It was actually what first got me into cave exploring, with its hehnia formations that are unique to this region. So I was excited to return to it to try to map it.

We met with Mr. Kloht, a local resident who knows the cave well and who would be joining me. He gave me a semev to use on the mission, which is a kind of echolocation tool he uses to gauge the distance of a chasm.

The first day was just me and Mr. Kloht. After going through some long, winding tunnels, the cave opened up to a pretty open cavern. We spent time taking pictures of some of the different formations we observed. There was one particularly large hehnia formation in the center of the cavern. That helped serve as a nice anchoring point when we then worked on taking measurements for accurate scaling of the map.

Before heading back out at the end of the day, we made sure to note all of the different tunnels and chasms possibly leading to other areas. There was one dauntingly narrow tunnel. And when we used the semev to gauge it, it seemed like it was quite long as well. It was too narrow for either of us to fit into, but Allison might. One thing was sure: There was a lot more of Slaqsi cave remaining to be explored in the weeks to come.

### **Trial 1 (“Exploring”) sequel, Iraq framing:**

Allison was intrigued to hear about how the first day in Slaqsi cave had gone. She had read about the hehnia formations and wanted a chance to see them in person. These were probably her biggest reason for wanting to go on this mission.

When Mr. Kloht described how long the initial tunnels were, she thought that sounded kind of boring. But she was excited to hear about the really narrow opening they had found. They had used the semev for an estimation, suggesting that the tunnel was quite long. Allison had become quite an expert at narrow passages, so she jumped at the chance when they said they’d need her to do that one solo.

Then they reviewed their measurements and photos to start creating a map. Slaqsi cave definitely photographed well. The hehnia pictures they had taken looked amazing. She was excited to see them for herself. And the measurements from the main cavern were surprisingly large.

At the end of the day, Allison was glad to have joined the mission. She also had lots of prep to do. She would need to use the semev for her solo mission, so she’d make sure to get Mr. Kloht to teach her how. And she’d need to make sure she had the best gear for tight spaces.

### **Trial 1 (“Exploring”) exposure story, Indonesia framing:**

Last year, I was the lead on the biggest mission my cave exploration team had ever taken on. We would be the first team to officially map Slaksi cave, located in Western Indonesia. Allison was my assistant who traveled with me and helped me prepare, and I translated between her and the local residents.

My hometown being just an hour’s drive away, I had been to the cave once before when I was a kid. It was actually what first got me into cave exploring, with its hénia formations that are unique to this region. So I was excited to return to it to try to map it.

We met with Mr. Klót, a local resident who knows the cave well and who would be joining me. He gave me a semév to use on the mission, which is a kind of echolocation tool he uses to gauge the distance of a chasm.

The first day was just me and Mr. Klót. After going through some long, winding tunnels, the cave opened up to a pretty open cavern. We spent time taking pictures of some of the different formations we observed. There was one particularly large hénia formation in the center of the cavern. That helped serve as a nice anchoring point when we then worked on taking measurements for accurate scaling of the map.

Before heading back out at the end of the day, we made sure to note all of the different tunnels and chasms possibly leading to other areas. There was one dauntingly narrow tunnel. And when we used the semév to gauge it, it seemed like it was quite long as well. It was too narrow for either of us to fit into, but Allison might. One thing was sure: There was a lot more of Slaksi cave remaining to be explored in the weeks to come.

### **Trial 1 (“Exploring”) sequel, Indonesia framing:**

Allison was intrigued to hear about how the first day in Slaksi cave had gone. She had read about the hénia formations and wanted a chance to see them in person. These were probably her biggest reason for wanting to go on this mission.

When Mr. Klót described how long the initial tunnels were, she thought that sounded kind of boring. But she was excited to hear about the really narrow opening they had found. They had used the semév for an estimation, suggesting that the tunnel was quite long. Allison had become quite an expert at narrow passages, so she jumped at the chance when they said they’d need her to do that one solo.

Then they reviewed their measurements and photos to start creating a map. Slaksi cave definitely photographed well. The hénia pictures they had taken looked amazing. She was excited to see them for herself. And the measurements from the main cavern were surprisingly large.

At the end of the day, Allison was glad to have joined the mission. She also had lots of prep to do. She would need to use the semév for her solo mission, so she’d make sure to get Mr. Klót to teach her how. And she’d need to make sure she had the best gear for tight spaces.

## **Trial 2 (“Wedding”) exposure story, US framing:**

My wife is from California. But we had our wedding in the small town of Sondow in the north of Maine that I’m from (and that most no one has heard of). Her family, the Millers, came all the way from LA. Fortunately they seemed quite charmed with the place.

When they arrived, my parents had made sheenia for dinner, which is a fish that’s pretty much the only thing Sondow is known for. My wife’s parents really liked it, though my sister-in-law-to-be was not super thrilled with it.

After dinner, we took them to the local park where the town troubadour was playing a zenool flute, named after a special reed he grows and carves them out of. The family was impressed, saying he was much better than any street performer they’d seen in LA.

The next day we took them to the wedding venue. Mrs. Miller really loved the place. She especially loved the cume flowers, which she’d never seen before. She asked if she could have one to wear in her hair on the wedding day. My sister-in-law-to-be was distracted in making sure that sheenia wasn’t going to be the only option for dinner at the reception. I assured her that it wouldn’t be.

The night before the wedding we all had dinner together again: both families. We made sure the table was decorated with a lot of cume flowers to please my mother-in-law. After we saw the town troubadour the other night, I went back to buy a zenool flute for each of them to give as presents at the dinner. I think they liked them; but in hindsight, I doubt that they’ve ever tried to play them since.

## **Trial 2 (“Wedding”) sequel, US framing:**

The Millers had such a nice time at the wedding. The small, charming town of Sondow was a nice change from busy LA.

The wedding was beautiful. There were a couple small hiccups, like most weddings, but nothing major. The cume flower fell out of Mrs. Miller’s hair right before going down the aisle, but one of the bridesmaids helped her put it back in. The bride’s sister was given sheenia for dinner by accident at the reception. Fortunately, getting the caterers to serve her something else didn’t create a scene.

The Millers are back in LA now. The zenool flutes they received are resting on their mantle as a reminder of the good memories they made. They sit beside lots of pictures from the Millers’ trip and the wedding: pictures of the troubadour playing, the wedding ceremony, and the venue with all the cume flowers everywhere.

They sent some gifts to their new son-in-law’s parents back in Sondow. And sheenia filet is a new craving that Mr. Miller experiences often. He was sad that he didn’t get to go fishing while he was there to try to catch one himself. The family has decided to try to save money to go back, though. And their new extended family told them they would definitely have a place to stay.

The zenool flute CD they bought from the troubadour plays in the background. Mrs. Miller tried to learn to play the instrument, herself. But she had never had much luck in previous musical endeavors. This time was no different.

## **Trial 2 (“Wedding”) exposure story, Iraq framing:**

My wife is from California. But we had our wedding in the small town of Sandoh in Iraq that I’m from (and that most no one has heard of). Her family, the Millers, came all the way from LA, and I was worried about them not speaking the language. Fortunately they seemed quite charmed with the place.

When they arrived, my parents had made shinia for dinner, which is a fish that’s pretty much the only thing Sandoh is known for. My wife’s parents really liked it, though my sister-in-law-to-be was not super thrilled with it.

After dinner, we took them to the local park where the town troubadour was playing a zenuhl flute, named after a special reed he grows and carves them out of. The family was impressed, saying he was much better than any street performer they’d seen in LA.

The next day we took them to the wedding venue. Mrs. Miller really loved the place. She especially loved the kjum flowers, which she’d never seen before. She asked if she could have one to wear in her hair on the wedding day. My sister-in-law-to-be was distracted in making sure that shinia wasn’t going to be the only option for dinner at the reception. I assured her that it wouldn’t be.

The night before the wedding we all had dinner together again: both families. We made sure the table was decorated with a lot of kjum flowers to please my mother-in-law. After we saw the town troubadour the other night, I went back to buy a zenuhl flute for each of them to give as presents at the dinner. I think they liked them; but in hindsight, I doubt that they’ve ever tried to play them since.

## **Trial 2 (“Wedding”) sequel, Iraq framing:**

The Millers had such a nice time at the wedding. The small, charming town of Sandoh was a nice change from busy LA.

The wedding was beautiful. There were a couple small hiccups, like most weddings, but nothing major. The kjum flower fell out of Mrs. Miller’s hair right before going down the aisle, but one of the bridesmaids helped her put it back in. The bride’s sister was given shinia for dinner by accident at the reception. Fortunately, getting the caterers to serve her something else didn’t create a scene.

The Millers are back in LA now. The zenuhl flutes they received are resting on their mantle as a reminder of the good memories they made. They sit beside lots of pictures from the Millers’ trip and the wedding: pictures of the troubadour playing, the wedding ceremony, and the venue with all the kjum flowers everywhere.

They sent some gifts to their new son-in-law’s parents back in Sandoh. And shinia filet is a new craving that Mr. Miller experiences often. He was sad that he didn’t get to go fishing while he was there to try to catch one himself. The family has decided to try to save money to go back, though. And their new extended family told them they would definitely have a place to stay.

The zenuhl flute CD they bought from the troubadour plays in the background. Mrs. Miller tried to learn to play the instrument, herself. But she had never had much luck in previous musical endeavors. This time was no different.

## **Trial 2 (“Wedding”) exposure story, Indonesia framing:**

My wife is from California. But we had our wedding in the small town of Sándo in Indonesia that I’m from (and that most no one has heard of). Her family, the Millers, came all the way from LA, and I was worried about them not speaking the language. Fortunately they seemed quite charmed with the place.

When they arrived, my parents had made shínia for dinner, which is a fish that’s pretty much the only thing Sándo is known for. My wife’s parents really liked it, though my sister-in-law-to-be was not super thrilled with it.

After dinner, we took them to the local park where the town troubadour was playing a zenúl flute, named after a special reed he grows and carves them out of. The family was impressed, saying he was much better than any street performer they’d seen in LA.

The next day we took them to the wedding venue. Mrs. Miller really loved the place. She especially loved the kyum flowers, which she’d never seen before. She asked if she could have one to wear in her hair on the wedding day. My sister-in-law-to-be was distracted in making sure that shínia wasn’t going to be the only option for dinner at the reception. I assured her that it wouldn’t be.

The night before the wedding we all had dinner together again: both families. We made sure the table was decorated with a lot of kyum flowers to please my mother-in-law. After we saw the town troubadour the other night, I went back to buy a zenúl flute for each of them to give as presents at the dinner. I think they liked them; but in hindsight, I doubt that they’ve ever tried to play them since.

## **Trial 2 (“Wedding”) sequel, Indonesia framing:**

The Millers had such a nice time at the wedding. The small, charming town of Sándo was a nice change from busy LA.

The wedding was beautiful. There were a couple small hiccups, like most weddings, but nothing major. The kyum flower fell out of Mrs. Miller’s hair right before going down the aisle, but one of the bridesmaids helped her put it back in. The bride’s sister was given shínia for dinner by accident at the reception. Fortunately, getting the caterers to serve her something else didn’t create a scene.

The Millers are back in LA now. The zenúl flutes they received are resting on their mantle as a reminder of the good memories they made. They sit beside lots of pictures from the Millers’ trip and the wedding: pictures of the troubadour playing, the wedding ceremony, and the venue with all the kyum flowers everywhere.

They sent some gifts to their new son-in-law’s parents back in Sándo. And shínia filet is a new craving that Mr. Miller experiences often. He was sad that he didn’t get to go fishing while he was there to try to catch one himself. The family has decided to try to save money to go back, though. And their new extended family told them they would definitely have a place to stay.

The zenúl flute CD they bought from the troubadour plays in the background. Mrs. Miller tried to learn to play the instrument, herself. But she had never had much luck in previous musical endeavors. This time was no different.

### **Trial 3 (“Visiting”) exposure story, US framing:**

Heading home for spring vacation that year, I was excited that my good friend and roommate John was going to come with me. John’s family lives in Nebraska, and he had never been to South Carolina—where I’m from.

My dad picked the two of us up from the airport. While he drove, he was asking John about his hobbies and his family. When we arrived at the house, Mrs. Glay, our neighbor and good family friend, was outside gardening and waved to us. John remarked that the trees in front were very interesting. I told him that they’re deneer trees, which he had never heard of before. I explained how their intertwining trunks were perfect for climbing and reminisced climbing them all the time when I was young.

After we unpacked and ate lunch, we went for a walk to show John the beach that’s close by. We ran into Mrs. Glay, who introduced herself. Learning that John was new to the area, she pointed out a noomia to him, which is basically a seagull but with reversed colors. She then proceeded to talk for several minutes about how much of a nuisance she finds them.

We parted ways and kept walking to the beach. I told him that we should try to find him a gurgee shell, which he’d probably never seen before. When we got there, John saw that there are a lot of deneer trees lining the sand. So he insisted that we take some time to do some tree climbing. He definitely enjoyed himself and agreed that these were the best tree he’d ever climbed.

Then we swam for a while and had fun combing the beach for brightly colored rocks and oddly shaped shells. I found a perfect gurgee shell that I gave to John. Just as he was admiring it, a noomia swept right by him, making him jolt and cast it far away by accident. Fortunately, we found it again in the sand and had a good laugh.

The week went by surprisingly fast, but we had a lot of fun. He seemed to have had a great time meeting my family and seeing my home.

### **Trial 3 (“Visiting”) sequel, US framing:**

John had a great time visiting South Carolina for the first time with his roommate. Back at school, he researched whether a deneer tree planted farther north would grow. He was sad to learn that it wouldn’t. The two roommates decided that he would just have to come back sometime to admire them again.

John also said that he wanted to find more gurgee shells. He recalled when the noomia flew past him, making him lose his shell. The two laughed reminiscently. John said that he can now completely sympathize with Mrs. Glay’s disdain for them: seeing one perched near him would make him a little angry now, too.

Mrs. Glay mailed John a picture she had taken of him. He was climbing the deneer tree in front of the house and had gotten quite high up. He looked very pleased with himself. The picture was just a split second before he then lost his footing and almost fell all the way down.

John felt a lot closer to his roommate, getting to see where he grew up and meet his family. He felt like he had a new community there, himself. He and his roommate had made a lot of good memories together. The story of the noomia flying at John gets recounted often to friends and family. John also now keeps the gurgee shell he got on his desk as a decoration.

### **Trial 3 (“Visiting”) exposure story, Iraq framing:**

Heading home for spring vacation that year, I was excited that my good friend and roommate John was going to come with me. John’s family lives in Nebraska, and he had never been to Iraq—where I’m from.

My dad picked the two of us up from the airport. While he drove, he was asking John about his hobbies and his family. (I would step in to translate every now and then when John had a hard time understanding my dad’s English.) When we arrived at the house, Mrs. Gleh, our neighbor and good family friend, was outside gardening and waved to us. John remarked that the trees in front were very interesting. I told him that they’re denihr trees, which he had never heard of before. I explained how their intertwining trunks were perfect for climbing and reminisced climbing them all the time when I was young.

After we unpacked and ate lunch, we went for a walk to show John the beach that’s close by. We ran into Mrs. Gleh, who introduced herself. Learning that John was new to the area, she pointed out a numia to him, which is basically a seagull but with reversed colors. She then proceeded to talk for several minutes about how much of a nuisance she finds them.

We parted ways and kept walking to the beach. I told him that we should try to find him a gurdzhi shell, which he’d probably never seen before. When we got there, John saw that there are a lot of denihr trees lining the sand. So he insisted that we take some time to do some tree climbing. He definitely enjoyed himself and agreed that these were the best tree he’d ever climbed.

Then we swam for a while and had fun combing the beach for brightly colored rocks and oddly shaped shells. I found a perfect gurdzhi shell that I gave to John. Just as he was admiring it, a numia swept right by him, making him jolt and cast it far away by accident. Fortunately, we found it again in the sand and had a good laugh.

The week went by surprisingly fast, but we had a lot of fun. He seemed to have had a great time meeting my family and seeing my home.

### **Trial 3 (“Visiting”) sequel, Iraq framing:**

John had a great time visiting Iraq for the first time with his roommate. Back at school, he researched whether a denihr tree planted in the US would grow. He was sad to learn that it wouldn’t. The two roommates decided that he would just have to come back sometime to admire them again.

John also said that he wanted to find more gurdzhi shells. John recalled when the numia flew past him, making him lose his shell. The two laughed reminiscently. John said that he can now completely sympathize with Mrs. Gleh’s disdain for them: seeing one perched near him would make him a little angry now, too.

Mrs. Gleh mailed John a picture she had taken of him. He was climbing the denihr tree in front of the house and had gotten quite high up. He looked very pleased with himself. The picture was just a split second before he then lost his footing and almost fell all the way down.

John felt a lot closer to his roommate, getting to see where he grew up and meet his family. He felt like he had a new community there, himself. He and his roommate had made a lot of good memories together. The story of the numia flying at John gets recounted often to friends and family. John also now keeps the gurdzhi shell he got on his desk as a decoration.



### **Trial 3 (“Visiting”) exposure story, Indonesia framing:**

Heading home for spring vacation that year, I was excited that my good friend and roommate John was going to come with me. John’s family lives in Nebraska, and he had never been to Indonesia—where I’m from.

My dad picked the two of us up from the airport. While he drove, he was asking John about his hobbies and his family. (I would step in to translate every now and then when John had a hard time understanding my dad’s English.) When we arrived at the house, Mrs. Glé, our neighbor and good family friend, was outside gardening and waved to us. John remarked that the trees in front were very interesting. I told him that they’re denír trees, which he had never heard of before. I explained how their intertwining trunks were perfect for climbing and reminisced climbing them all the time when I was young.

After we unpacked and ate lunch, we went for a walk to show John the beach that’s close by. We ran into Mrs. Glé, who introduced herself. Learning that John was new to the area, she pointed out a númera to him, which is basically a seagull but with reversed colors. She then proceeded to talk for several minutes about how much of a nuisance she finds them.

We parted ways and kept walking to the beach. I told him that we should try to find him a gurji shell, which he’d probably never seen before. When we got there, John saw that there are a lot of denír trees lining the sand. So he insisted that we take some time to do some tree climbing. He definitely enjoyed himself and agreed that these were the best tree he’d ever climbed.

Then we swam for a while and had fun combing the beach for brightly colored rocks and oddly shaped shells. I found a perfect gurji shell that I gave to John. Just as he was admiring it, a númera swept right by him, making him jolt and cast it far away by accident. Fortunately, we found it again in the sand and had a good laugh.

The week went by surprisingly fast, but we had a lot of fun. He seemed to have had a great time meeting my family and seeing my home.

### **Trial 3 (“Visiting”) sequel, Indonesia framing:**

John had a great time visiting Indonesia for the first time with his roommate. Back at school, he researched whether a denír tree planted in the US would grow. He was sad to learn that it wouldn’t. The two roommates decided that he would just have to come back sometime to admire them again.

John also said that he wanted to find more gurji shells. John recalled when the númera flew past him, making him lose his shell. The two laughed reminiscently. John said that he can now completely sympathize with Mrs. Glé’s disdain for them: seeing one perched near him would make him a little angry now, too.

Mrs. Glé mailed John a picture she had taken of him. He was climbing the denír tree in front of the house and had gotten quite high up. He looked very pleased with himself. The picture was just a split second before he then lost his footing and almost fell all the way down.

John felt a lot closer to his roommate, getting to see where he grew up and meet his family. He felt like he had a new community there, himself. He and his roommate had made a lot of good memories together. The story of the númera flying at John gets recounted often to friends and family. John also now keeps the gurji shell he got on his desk as a decoration.

## G Step-up modeling, Nonce word imitation

### Categorical imitation:

#### Base:

```
lmer(avg.num.of.divergences ~ framing
      + (1|word) + (1|word:framing),
      data = data)
```

#### Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
(intercept)	.164	.093	1.755	
framing.Indonesia	.297	.089	3.302	.00097
framing.Iraq	.339	.089	3.773	

\*Pr(>Chisq) comes from Chi-square ANOVA test between models including and excluding the fixed effect of interest.

#### Step 1 (lone-predictor models):

```
lmer(avg.num.of.divergences ~ fixed.effect
      + framing
      + (1|word) + (1|word:framing),
      data = data)
```

#### Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
globalist/nationalist alignment	.042	.014	3.036	.0024
Iraq-directed attitude (Likert)**	-.004	.024	-.015	.01
political identity**	.036	.024	1.49	.02
mobility	-.018	.014	-1.288	.195
language contact ideology	.018	.014	1.281	.197
prescriptivism	-.005	.014	-.338	.738
Iraq-directed attitude (IAT)	.002	.014	.131	.904

\*Pr(>Chisq) comes from Chi-square ANOVA test between model from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ fixed.effect:framing” term was compared to a model without, to test if including an interaction between the fixed effect of interest and framing was motivated. For this fixed effect, the inclusion of an interaction term with framing was found to significantly improve model fit.

#### Step 2:

```
lmer(avg.num.of.divergences ~ fixed.effect
      + globalist.nationalist.alignment
      + framing)
```

```
+ (1|word) + (1|word : framing),
data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
Iraq-directed attitude (Likert)**	-.008	.024	-.339	.019
political identity**	.023	.025	.932	.044
mobility	-.024	.014	-1.667	.094
prescriptivism	.01	.015	.692	.485
Iraq-directed attitude (IAT)	.004	.014	.309	.765
language contact ideology	-.005	.016	-.299	.769

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ fixed.effect : framing” term was compared to a model without, to test if including an interaction between the fixed effect of interest and framing was motivated. For this fixed effect, the inclusion of an interaction term with framing was found to significantly improve model fit.

Step 3:

```
lmer(avg.num.of.divergences ~ fixed.effect
+ globalist.nationalist.alignment + Iraq.Likert
+ framing
+ (1|word) + (1|word : framing),
data = data)
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
political identity**	.025	.025	.998	.053
mobility	-.018	.015	-1.213	.221
prescriptivism	.015	.015	1.033	.298
language contact ideology	-.015	.017	-.896	.373
Iraq-directed attitude (IAT)	.001	.014	.083	.94

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ fixed.effect : framing” term was compared to a model without, to test if including an interaction between the fixed effect of interest and framing was motivated. For this fixed effect, the inclusion of an interaction term with framing was found to significantly improve model fit.

Final model:

```
lmer(avg.num.of.divergences ~
globalist.nationalist.alignment + Iraq.Likert
+ framing
+ (1|word) + (1|word : framing),
```

```
data = data)
```

Results:

```
("intercept" = US framing)
```

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.414	-0.476	-0.162	0.318	4.781

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
(intercept)	.163	.093	1.754	
framing.Indonesia	.297	.089	3.352	.00091
framing.Iraq	.339	.089	3.834	
globalist/nationalist alignment	.038	.015	2.54	.011
Iraq.Likert : framing.US	-.009	.024	-0.404	.023
Iraq.Likert : framing.Indonesia	.071	.024	2.901	
Iraq.Likert : framing.Iraq	.023	.024	0.942	
political.identity : framing.US	.025	.025	0.998	.053
political.identity : framing.Indonesia	.039	.025	1.594	
political.identity : framing.Iraq	-.049	.025	-1.969	

\*Pr(>Chisq) comes from Chi-square ANOVA test between full model and model excluding component of interest.

**Phonetic imitation, vowel reduction variable:**

Base:

```
lmer(heard.likeness ~ (1|exposure.variant),  
data = data[data$variable == "vowel.reduction",])
```

Results:

Fixed effect	Estimate	Std. Err.	t value
(intercept)	.51	.085	5.968

Step 1 (lone-predictor models):

```
lmer(heard.likeness ~ fixed.effect  
+ (1|exposure.variant),  
data = data[data$variable == "vowel.reduction",])
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
Iraq-directed attitude (IAT)	-.009	.009	-1.037	.297
Iraq-directed attitude (Likert)	-.008	.009	-.777	.432
language contact ideology	-.006	.009	-.629	.523
mobility	.006	.009	.608	.54
political identity	.002	.009	.217	.83
globalist/nationalist alignment	.001	.009	.107	.914
prescriptivism	-.0001	.009	-.011	.994

\*Pr(>Chisq) comes from Chi-square ANOVA test between model from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ fixed.effect:framing” term was compared to a model without, to test if including an interaction between the fixed effect of interest and framing was motivated. For this fixed effect, the inclusion of an interaction term with framing was found to significantly improve model fit.

### Phonetic imitation, gliding variable:

Base:

```
lmer(heard.likeness ~ (1|exposure.variant),
     data = data[data$variable == ‘gliding’,])
```

Results:

Fixed effect	Estimate	Std. Err.	t value
(intercept)	.562	.089	6.282

Step 1 (lone-predictor models):

```
lmer(heard.likeness ~ fixed.effect
     + (1|exposure.variant),
     data = data[data$variable == ‘gliding’,])
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
Iraq-directed attitude (IAT)**	.041	.029	1.404	.023
globalist/nationalist alignment	-.037	.017	-2.161	.029
prescriptivism	-.019	.017	-1.143	.251
Iraq-directed attitude (Likert)	-.012	.017	-0.722	.47
mobility	.005	.017	0.288	.777
political identity	-.004	.017	-0.243	.808
language contact ideology	-.003	.017	-0.166	.87

\*Pr(>Chisq) comes from Chi-square ANOVA test between model from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ fixed.effect : framing” term was compared to a model without, to test if including an interaction between the fixed effect of interest and framing was motivated. For this fixed effect, the inclusion of an interaction term with framing was found to significantly improve model fit.

### Step 2:

```
lmer(heard.likeness ~ fixed.effect + Iraq.IAT
      + (1|exposure.variant),
      data = data[data$variable == ‘gliding’,])
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
globalist/nationalist alignment	-.035	.017	-2.114	.032
prescriptivism	-.019	.017	-1.163	.239
Iraq-directed attitude (Likert)	.552	-.013	-.784	.43
mobility	.004	.017	.232	.819
political identity	-.003	.017	-0.195	.845
language contact ideology	-.0007	.017	-0.039	.971

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ fixed.effect : framing” term was compared to a model without, to test if including an interaction between the fixed effect of interest and framing was motivated. For this fixed effect, the inclusion of an interaction term with framing was found to significantly improve model fit.

### Step 3:

```
lmer(heard.likeness ~ fixed.effect
      + Iraq.IAT + globalist.nationalist.alignment
      + (1|exposure.variant),
      data = data[data$variable == ‘gliding’,])
```

Results:

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
prescriptivism	-.019	.017	-1.163	.239
Iraq-directed attitude (Likert)	-.013	.017	-0.784	.43
mobility	.004	.017	0.232	.819
political identity	-.003	.017	-0.195	.845
language contact ideology	-.0007	.017	-0.039	.971

\*Pr(>Chisq) comes from Chi-square ANOVA test between model containing top fixed effect from previous step and an identical model additionally including the fixed effect of interest in the current step.

\*\*For each fixed effect, a model including a “+ fixed.effect : framing” term was compared to a model without, to test if including an interaction between the fixed effect of interest and framing was motivated. For this fixed effect, the inclusion of an interaction term with framing was found to significantly improve model fit.

Final model:

```
lmer(heard.likeness ~  
      Iraq.IAT:framing + globalist.nationalist.alignment  
      + (1|exposure.variant),  
      data = data[data$variable == "gliding",])
```

Results:

("intercept" = US framing)

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.442	-0.791	0.126	0.803	1.885

Fixed effect	Estimate	Std. Err.	t value	Pr(>Chisq)
(intercept)	.561	.083	6.785	
globalist/nationalist alignment	-.036	.017	-2.131	.031
Iraq.IAT:framing.US	.038	.029	1.344	.025
Iraq.IAT:framing.Indonesia	.056	.029	1.93	
Iraq.IAT:framing.Iraq	-.056	.029	-1.94	

\*Pr(>Chisq) comes from Chi-square ANOVA test between full model and model excluding component of interest.

## References

- AboveTopSecret.com (2009). Does anyone else hate the way obama pronounces pakistan. URL <http://www.abovetopsecret.com/forum/thread523647/pg1>
- Abramowitz, A. (2013). The electoral roots of America's dysfunctional government. *Presidential studies quarterly*, 43(4):709–731. doi:10.1111/psq.12063
- Abramowitz, A. and Saunders, K. (1998). Ideological realignment in the US electorate. *Journal of politics*, 60:634–52. doi:10.2307/2647642
- Adler, A. (2006). Faithfulness and perception in loanword adaptation: A case study from Hawaiian. *Lingua*, 116:1024–1045. doi:10.1016/j.lingua.2005.06.007
- Alfaraz, G. (2002). Miami Cuban perceptions of varieties of Spanish. In Long, D. and Preston, D., editors, *Handbook of perceptual dialectology*, volume 2, pages 1–12. Amsterdam: John Benjamins.
- Anderson, R., Klostad, C., Mayew, W., and Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLoS ONE*, 9(5):e97506. doi:10.1371/journal.pone.0097506
- "Antonio Mendoza" (1990, November 10). *Saturday Night Live [Television broadcast]*. New York: WNBC.
- Arab American Institute (2014). American attitudes toward Arabs and Muslims. URL <http://www.aaiusa.org/american-attitudes-toward-arabs-and-muslims-2014>
- Arsenault, P. (2009). Coronal features and retroflexion in Dhivehi and other Indo-Aryan languages. *Toronto working papers in linguistics*, 30:17–33.



- Babel, M. (2009). *Phonetic and social selectivity in speech accommodation*. PhD thesis, University of California, Berkeley.
- Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in society*, 39(4):437–456. doi:10.1017/S0047404510000400
- Bailey, G. (2002). Real and apparent time. In Chambers, J., Trudgill, P., and Schilling-Estes, N., editors, *The handbook of language variation and change*, pages 312–332. Oxford: Blackwell.
- Baker, M. (2016, December 5). Norwegian Air plans U.S. expansion. *Business Travel News*. URL <http://www.businesstravelnews.com/Global/Norwegian-Air-Plans-US-Expansion>
- Baker, W., Eddington, D., and Nay, L. (2009). Dialect identification: The effects of region of origin and amount of experience. *American speech*, 84(1):48–71. doi:10.1215/00031283-2009-004
- Banikowski, B. and DiMaggio, P. (2016). Varieties of American popular nationalism. *American sociological review*, 81(5):949–980. doi:10.1177/0003122416663683
- Baran, D. (2017). *Language in Immigrant America*. Cambridge: Cambridge University Press. doi:10.1017/9781107415713
- Barker, V., Giles, H., Noels, K., Duck, J., Hecht, M., and Clément, R. (2001). The English-only movement: A communication analysis of changing perceptions of language vitality. *International Communication Association*, 51:3–37. doi:10.1111/j.1460-2466.2001.tb02870.x
- Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.
- Bauer, L. and Trudgill, P., editors (1998). *Language myths*. New York: Penguin Books.

- Baugh, A. and Cable, T. (2013). *A history of the English language*. London: Pearson, 6 edition.
- Beckwith, R. (2016, March 22). Read President Obama's speech to the Cuban people. *Time*. URL <http://time.com/4267933/barack-obama-cuba-speech-transcript-full-text/>
- Bell, A. (1984). Language style as audience design. *Language in society*, 13(2):145–204. doi:10.1017/S004740450001037X
- Bergen, B. (2004). The psychological reality of phonaesthemes. *Language*, 80(2):290–311. doi:10.1353/lan.2004.0056
- Blake, R. (2016). Toward heterogeneity: A sociolinguistic perspective on the classification of Black people in the twenty-first century. In Alim, H. S., Rickford, J., and Ball, A., editors, *Raciolinguistics: How language shapes our ideas about race*, pages 153–69. Oxford: Oxford University Press.
- Block, D., Gray, J., and Holborow, M. (2012). *Neoliberalism and applied linguistics*. New York: Routledge.
- Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.
- Boberg, C. (1997). *Variation and change in the nativization of foreign (a) in English*. PhD thesis, University of Pennsylvania.
- Boberg, C. (1999). The attitudinal component of variation in American English foreign (a) nativization. *Journal of language and social psychology*, 18(1):49–61. doi:10.1177/0261927X99018001004

- Boberg, C. (2014). Cross-border convergence in loan phonology: Foreign (a) in North American English. *Paper presented at New Ways of Analyzing Variation (NWAV) 43*, University of Illinois at Urbana-Champaign.
- Boersma, P. and Hamann, S. (2009). Loanword adaptation as first-language phonological perception. In Calabrese, A. and Wetzels, W. L., editors, *Loanw phonology*, pages 11–58. Amsterdam: John Benjamins.
- Boersma, P. and Weenink, D. (2015). Praat: Doing phonetics by computer [computer program], version 5.3.77. URL <http://www.fon.hum.uva.nl/praat/>
- Bourdieu, P. (1977). The economics of linguistic exchanges. *Social science information*, 16(6):645–68.
- Bourhis, R. and Giles, H. (1977). The language of intergroup distinctiveness. In Giles, H., editor, *Language, ethnicity and intergroup relations*, pages 119–135. London: Academic Press.
- Brannen, K. (2002). The role of perception in differential substitution. *Canadian journal of linguistics*, 47:1–46. doi:10.1017/S0008413100018004
- Broselow, E. (2009). Stress adaptation in loanword phonology: Perception and learnability. In Boersma, P. and Hamann, S., editors, *Phonology in perception*, pages 191–234. Berlin: De Gruyter.
- Bucholtz, M. (2016). On being called out of one’s name: Indexical bleaching as a technique of deracialization. In Alim, H. S., Rickford, J., and Ball, A., editors, *Raciolinguistics: How language shapes our ideas about race*, pages 273–289. Oxford: Oxford University Press.
- Bucholtz, M. and Hall, K. (2004). Language and identity. In Duranti, A., editor, *A companion to linguistic anthropology*, pages 369–394. Oxford: Blackwell.

- Callesano, S. and Carter, P. (2016). Oral vs. visual stimuli in the IAT: The case of Spanish and English in Miami. *Poster presented at New Ways of Analyzing Variation (NWAV) 45*, Simon Fraser University.
- Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American speech*, 82(1):32–64. doi:10.1215/00031283-2007-002
- Campbell-Kibler, K. (2010). Sociolinguistics and perception. *Language and linguistics compass*, 4(6):377–89. doi:10.1111/j.1749-818x.2010.00201.x
- Campbell-Kibler, K. (2012). The Implicit Association Test and sociolinguistic meaning. *Lingua*, 122(7):753–763. doi:10.1016/j.lingua.2012.01.002
- Caprara, G., Barbaranelli, C., Consiglio, C., Picconi, L., and Zimbardo, P. (2003). Personalities of politicians and voters: Unique and synergistic relationships. *Journal of personality and social psychology*, 84(4):849–856. doi:10.1037/0022-3514.84.4.849
- Caprara, G., Schwartz, S., Capanna, C., Vecchione, M., and Barbaranelli, C. (2006). Personality and politics: Values, traits, and political choice. *Political psychology*, 27(1):1–28. doi:10.1111/j.1467-9221.2006.00447.x
- Cashin, S. (2010). To be Muslim or Muslim-looking in America: A comparative exploration of racial and religious prejudice in the 21st century. *Duke forum for law and social change*, 2:125–139.
- Cavalli-Sforza, L., Piazza, A., Menozzi, P., and Mountain, J. (1988). Reconstruction of human evolution: Bringing together genetic, archaeological, and linguistic data. *Proceedings of the National Academy of Sciences*, 85(16):6002–06. doi:10.1073/pnas.85.16.6002
- Chang, C. (2012). Phonetics vs. phonology in loanword adaptation: Revisiting the role of the bilingual. In Berson, S., Bratkievich, A., Bruhn, D., Campbell, A., Escamilla, R., Giovine, A., Newbold, L., Perez, M., Piqueras-Brunet, M., and Rhomieux, R., editors, *Proceedings*

- of the 34th annual meeting of the Berkeley Linguistics Society, pages 61–72. Berkeley, CA: Berkeley Linguistics Society.
- Cheshire, J. (2002). Sex and gender in variationist research. In Chambers, J., Trudgill, P., and Schilling-Estes, N., editors, *The handbook of language variation and change*, pages 423–43. Oxford: Blackwell.
- Chirumbolo, A. and Leone, L. (2010). Personality and politics: The role of the HEXACO model of personality in predicting ideology and voting. *Personality and individual differences*, 49(1):43–48. doi:10.1016/j.paid.2010.03.004
- Chitoran, I. (2002). A perception-production study of Romanian diphthongs and glide-vowel sequences. *Journal of the International Phonetic Association*, 32(2):203–222. doi:10.1017/S0025100302001044
- Choe, S. (2013, November 22). South Korea is expected to buy Lockheed fighter jets. *New York Times*. URL <https://www.nytimes.com/2013/11/23/business/international/south-korea-expected-to-buy-lockheed-fighter-jets.html>
- Citrin, J., Reingold, B., Walters, E., and Green, D. (1990). The "official English" movement and the symbolic politics of language in the United States. *Political research quarterly*, 43(3):535–559. doi:10.1177/106591299004300307
- Clements, G. (2001). Representational economy in constraint-based phonology. In Hall, T., editor, *Distinctive feature theory*, pages 71–146. Berlin: De Gruyter.
- Clopper, C. and Pisoni, D. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of phonetics*, 32(1):111–40. doi:10.1016/S0095-4470(03)00009-3
- Coffeed.com (2009). Tanzania pronunciation. URL <http://www.coffeed.com/viewtopic.php?f=19&t=2873>

- Cohen, G. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of personality and social psychology*, 85(5):808–822. doi:10.1037/0022-3514.85.5.808
- Cooper, R. (1975). Introduction to language attitudes II. *International journal of the sociology of language*, 6:5–9.
- Coupland, N., Williams, A., and Garrett, P. (1994). The social meanings of Welsh English: Teachers' stereotyped judgements. *Journal of multilingual and multicultural development*, 15(6):471–89. doi:10.1080/01434632.1994.9994585
- C-SPAN (2009, June 16). Obama-Lee Myung-bak joint news conference [video]. C-SPAN. URL <https://www.c-span.org/video/?287070-1/obama-lee-myung-bak-joint-news-conference&start=1570&stop=1728>
- Cukor-Avila, P. and Bailey, G. (2001). The effects of the race of the interviewer on sociolinguistic fieldwork. *Journal of sociolinguistics*, 5(2):252–270. doi:10.1111/1467-9481.00150
- Dailey-O'Cain, J. (2000). The sociolinguistic distribution of and attitudes toward focuser *like* and quotative *like*. *Journal of sociolinguistics*, 4(1):60–80. doi:10.1111/1467-9481.00103
- Davidson, L. (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology*, 24(2):261–286. doi:10.1017/S0952675707001200
- Davidson, L. and Noyer, R. (1997). Loan phonology in Huave: Nativization and the ranking of faithfulness constraints. *Proceedings of the West Coast Conference on Formal Linguistics*, 15:65–79.
- Davis, S., Tsujimura, N., and Tu, J. (2012). Toward a taxonomy of loanword prosody. *Catalan journal of linguistics*, 11:13–39. doi:10.5565/rev/catj1.8
- de Jong, K. and Cho, M.-H. (2012). Loanword phonology and perceptual mapping:

- Comparing two corpora of Korean contact with English. *Language*, 88(2):341–368. doi:10.1353/lan.2012.0035
- de la Garza, R. and DeSipio, L. (1994). Overview: The link between individuals and electoral institutions. In de la Garza, R., Menchaca, M., and DeSipio, L., editors, *Barrio ballots*, pages 1–41. Boulder, CO: Westview Press.
- de Saussure, F. (1916). *Course in general linguistics*. New York: McGraw-Hill.
- Dingemanse, M., Blasi, D., Lupyan, G., Christiansen, M., and Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in cognitive sciences*, 19(10):367–93. doi:10.1016/j.tics.2015.07.013
- Drager, K. (2018). *Experimental research methods in sociolinguistics*. London: Bloomsbury.
- Dresher, E. (2009). *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., and Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of experimental psychology: Human perception and performance*, 25(6):1568–1580. doi:10.1037/0096-1523.25.6.1568
- Dupoux, E., Parlato, E., Frota, S., Hirose, Y., and Peperkamp, S. (2011). Where do illusory vowels come from? *Journal of memory and language*, 64(3):199–210. doi:10.1016/j.jml.2010.12.004
- Eckert, P. (2000). *Linguistic variation as social practice*. Oxford: Blackwell.
- Eckert, P. (2004). Adolescent language. In Finegan, E. and Rickford, J., editors, *Language in the USA: Themes for the twenty-first century*, pages 361–374. Cambridge: Cambridge University Press.
- Eckert, P. (2008). Variation and the indexical field. *Journal of sociolinguistics*, 12(4):453–476. doi:10.1111/j.1467-9841.2008.00374.x

- Ferrara, K. and Bell, B. (1995). Variation and discourse function of constructed dialogue introducers: The case of *be + like*. *American speech*, 70(3):265–90. doi:10.2307/455900
- Fischer, M. and Massey, D. (2004). The ecology of racial discrimination. *City & community*, 3(3):221–41. doi:10.1111/j.1535-6841.2004.00079.x
- Fleischhacker, H. (2001). Cluster-dependent epenthesis asymmetries. *UCLA Working Papers in Linguistics*, 7.
- Flores, N. (2015). A tale of two visions: Hegemonic Whiteness and bilingual education. *Educational policy*, 30(1):13–38. doi:10.1177/0895904815616482
- Flores-Bayer, I. (2015). Sociolinguistic variation in practice: The strategic use of Chicano English sh~ch in a political election. *Paper presented at New Ways of Analyzing Variation (NWAV) 44*, University of Toronto.
- Forscher, P., Lai, C., Axt, J., Ebersole, C., Herman, M., Devine, P., and Nosek, B. (2018). A meta-analysis of change in implicit bias. *PsyArXiv*, (preprint: 21 May, 2018). doi:10.17605/OSF.IO/DV8TU
- Fowler, C., Brown, J., Sabadini, L., and Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of memory and language*, 49(3):396–413. doi:10.1016/S0749-596X(03)00072-X
- Fox, R., Flege, J., and Munro, M. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *The Journal of the Acoustical Society of America*, 97:2540–2551. doi:10.1121/1.411974
- Friesner, M. (2009). *The social and linguistic predictors of the outcomes of borrowing in the speech community of Montreal*. PhD thesis, University of Pennsylvania.
- Gal, S. and Irvine, J. (1995). The boundaries of languages and disciplines: How ideologies



- construct difference. *Social research: Defining the boundaries of social inquiry*, 62(4):967–1001.
- Garrett, P. (2004). Language contact and contact languages. In Duranti, A., editor, *A companion to linguistic anthropology*, pages 46–72. Oxford: Blackwell.
- Gentilucci, M. and Bernardis, P. (2007). Imitation during phoneme production. *Neuropsychologia*, 45(3):608–615. doi:10.1016/j.neuropsychologia.2006.04.004
- Giles, H. (1970). Evaluative reactions to accents. *Educational review*, 22(3):211–27. doi:10.1080/0013191700220301
- Giles, H., Coupland, N., and Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In Giles, H., Coupland, J., and Coupland, N., editors, *Contexts of accommodation*, pages 1–68. Cambridge: Cambridge University Press.
- Gingrich, N. (1995). English as the American language. In *To Renew America*, chapter 15. New York: HarperCollins.
- Gingrich, Newt (@newtingrich) (2016, April 27). "Washington elites mock Trump for mispronouncing Tanzania. They don't get it. He said the most important word correctly: America. He gets it." Tweet. 27 April 2016, 11:06 AM.
- Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, 105(2):251–279. doi:10.1037/0033-295X.105.2.251
- Gouskova, M. (2001). Falling sonority onsets, loanwords, and Syllable Contact. In Andronis, M., Ball, C., Elston, H., and Neuvel, S., editors, *Papers from the 37th meeting of the Chicago Linguistic Society*, pages 175–186. Chicago: Chicago Linguistic Society.
- Graff, D., Labov, W., and Harris, W. (1986). Testing listeners' reactions to phonological markers of ethnic identity: A new method for sociolinguistic research. In Sankoff, D., editor, *Diversity and diachrony*, pages 45–58. Amsterdam: John Benjamins.

- Gray, R., Bryant, D., and Greenhill, S. (2010). On the shape and fabric of human history. *Philosophical transactions of the Royal Society B*, 365:3923–33. doi:10.1098/rstb.2010.0162
- Greenwald, A., McGhee, D., and Schwartz, J. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of personality and social psychology*, 74(6):1464. doi:10.1037/0022-3514.74.6.1464
- Greenwald, A., Nosek, B., and Banaji, M. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of personality and social psychology*, 85(2):197–216. doi:10.1037/0022-3514.85.2.197
- Greenwald, A., Poehlman, T., Uhlmann, E., and Banaji, M. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of personality and social psychology*, 97(1):17–41. doi:10.1037/a0015575
- Hafez, O. (1996). Phonological and morphological integration of loanwords into Egyptian Arabic. *Égypte/Monde Arabe (Première série)*, 27-28:383–410. doi:10.4000/ema.1958
- Hall-Lew, L., Coppock, E., and Starr, R. (2010). Indexing political persuasion: Variation in the Iraq vowels. *American Speech*, 85(1):91–102. doi:10.1215/00031283-2010-004
- Hall-Lew, L., Starr, R., and Coppock, E. (2012). Style-Shifting in the U.S. Congress: The foreign (a) vowel in "Iraq(i)". In Campoy, J. and Espinosa, J., editors, *Style-Shifting in public: New perspectives on stylistic variation*, pages 45–63. Amsterdam: John Benjamins.
- Haspelmath, M. (2009). Lexical borrowing: Concepts and issues. In Haspelmath, M., editor, *Loanwords in the world's languages: A comparative handbook*, pages 35–54. Berlin: De Gruyter.
- Haugen, E. (1950). The analysis of linguistic borrowing. *Language*, 26(2):145–69. doi:10.2307/410058

- Haugen, E. (1953). *The Norwegian language in America: A study in bilingual behavior*. Philadelphia: University of Pennsylvania Press.
- Hay, J., Warren, P., and Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of phonetics*, 34(4):458–484. doi:10.1016/j.wocn.2005.10.001
- Herd, J. (2005). Loanword adaptation and the evaluation of similarity. *Toronto working papers in linguistics*, 24:65–116.
- Hill, J. (1995). Mock Spanish: A site for the indexical reproduction of racism in American English. *Language & Culture: Symposium 2*. URL <http://language-culture.binghamton.edu/symposia/2/part1/index.html>
- Hockett, C. (1960). The origin of speech. *Scientific American*, 203:88–111.
- Holden, K. (1976). Assimilation rates of borrowing and phonological productivity. *Language*, 52:131–47. doi:10.2307/413213
- Hsieh, F., Kenstowicz, M., and Mou, X. (2009). Mandarin adaptations of coda nasals in English loanwords. In Calabrese, A. and Wetzels, W., editors, *Loan phonology*, pages 131–154. Amsterdam: John Benjamins.
- Huetlin, J. (2017, May 25). Axis of reason: Obama meets Merkel in Berlin. Crowds go wildly nostalgic. *Daily Beast*. URL <https://www.thedailybeast.com/axis-of-reason-obama-meets-merkel-in-berlin-crowds-go-wildly-nostalgic>
- Huffington Post staff (2014, February 12). It turns out Barack Obama’s French is pretty bad. *Huffington Post*. URL [https://www.huffingtonpost.com/2014/02/12/obama-french\\_n\\_4775000.html](https://www.huffingtonpost.com/2014/02/12/obama-french_n_4775000.html)
- Hyman, L. (1970). The role of borrowing in the justification of phonological grammars. *Studies in African linguistics*, 1:1–48.

- Ito, C. (2014). Loanword accentuation in Yanbian Korean: A weighted-constraints analysis. *Natural language & linguistic theory*, 32(2):537–92. doi:10.1007/s11049-013-9211-y
- Ito, C., Kang, Y., and Kenstowicz, M. (2006). The adaptation of Japanese loanwords into Korean. *Studies in loanword phonology: MIT working papers in linguistics*, 52:65–104.
- Itô, J. and Mester, A. (1999). The phonological lexicon. In Tsujimura, N., editor, *Handbook of Japanese linguistics*, pages 62–100. Oxford: Blackwell.
- Jaffe, A. (1999). *Ideologies in action: Language politics on Corsica*. Berlin: De Gruyter.
- Jagers, Z. (2018). Evidence and characterization of a glide-vowel distinction in American English. *Laboratory phonology*, 9(1):3. doi:10.5334/labphon.36
- Janda, R., Joseph, B., and Jacobs, N. (1994). Systematic hyperforeignisms as maximally external evidence for linguistic rules. In Lima, S., Corrigan, R., and Iverson, G., editors, *The reality of linguistic rules*, pages 67–92. Amsterdam: John Benjamins.
- John, O., Donahue, E., and Kentle, R. (1991). *The Big Five Inventory—versions 4a and 54*. Berkeley, CA: Institute of Personality and Social Research, University of California, Berkeley.
- John, O., Naumann, L., and Soto, C. (2008). Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In John, O., Robins, R., and Pervin, L., editors, *Handbook of personality: Theory and research*, pages 114–58. New York: Guilford Press.
- Johnson, D. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and linguistics compass*, 3(1):359–383. doi:10.1111/j.1749-818X.2008.00108.x
- Jost, J., Glaser, J., Kruglanski, A., and Sulloway, F. (2003). Political conservatism as

- motivated social cognition. *Psychological bulletin*, 129(3):339. doi:10.1037/0033-2909.129.3.339
- Jost, J., Nosek, B., and Gosling, S. (2008). Ideology: Its resurgence in social, personality, and political psychology. *Perspectives on psychological science*, 3(2):126–136. doi:10.1111/j.1745-6916.2008.00070.x
- Kang, H. (1996). English loanwords in Korean. *Studies in phonetics, phonology, and morphology*, 2:21–47. Seoul: Hankwukmunhwasa.
- Kang, Y. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology*, 20(2):219–273. doi:10.1017/S0952675703004524
- Kang, Y. (2010). The emergence of phonological adaptation from phonetic adaptation: English loanwords in Korean. *Phonology*, 27(2):225–253. doi:10.1017/S0952675710000114
- Kang, Y. (2011). Loanword phonology. In van Oostendorp, M., Ewen, C. J., Hume, E., and Rice, K., editors, *The Blackwell companion to phonology*, volume 4, pages 2258–2282. Oxford: Blackwell.
- Kang, Y., Phạm, A., and Storme, B. (2016). French loanwords in Vietnamese: The role of input phonotactics and contrast in loanword adaptation. In Albright, A. and Fullwood, M., editors, *Proceedings of the 2014 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America. URL <https://journals.linguisticsociety.org/proceedings/index.php/amphonology/article/view/3749/3467>
- Karttunen, F. (1977). Finnish in America: A case study in monogenerational language change. In Blount, B. and Sanches, M., editors, *The social dimensions of language change*, pages 173–84. New York: Academic Press.
- Katayama, M. (1998). *Optimality Theory and Japanese loanword phonology*. PhD thesis, University of California, Santa Cruz.

- Kay, G. (1995). English loanwords in Japanese. *World Englishes*, 14(1):67–76. doi:10.1111/j.1467-971X.1995.tb00340.x
- Kenstowicz, M. (2005). The phonetics and phonology of Korean loanword adaptation. In Rhee, S., editor, *Proceedings of the First European Conference on Korean Linguistics*, pages 17–32. Seoul: Hankook Publishing.
- Kenstowicz, M. and Suchato, A. (2006). Issues in loanword adaptation: A case study from Thai. *Lingua*, 116(7):921–949. doi:10.1016/j.lingua.2005.05.006
- Kerswill, P. and Williams, A. (2002). Dialect recognition and speech community focusing in new and old towns in England: The effects of dialect levelling, demography and social networks. In Long, D. and Preston, D., editors, *Handbook of perceptual dialectology*, volume 2, pages 173–205. Amsterdam: John Benjamins.
- Kim, H. (2009). Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation. In Calabrese, A. and Wetzels, W., editors, *Loan phonology*, pages 155–80. Amsterdam: John Benjamins.
- Kim, H. (2017). A two-decade-interval variation in vowel insertion after word-final English and French postvocalic plosives in Korean adaptation: A sociolinguistic account. *Linguistics vanguard*, 3(1). doi:10.1515/lingvan-2016-0076
- Kiparsky, P. (1968). How abstract is phonology? In Fujimura, O., editor, *Three dimensions of linguistic theory*, pages 5–56. Tokyo: TEC.
- Kobach, K. (2016, September 1). *Fox & Friends* [Television broadcast]. *New York, NY: Fox*. URL [https://archive.org/details/FOXNEWSW\\_20160901\\_100000\\_FOX\\_\\_Friends/start/4740/end/4800](https://archive.org/details/FOXNEWSW_20160901_100000_FOX__Friends/start/4740/end/4800)
- Kohli, R. and Solórzano, D. (2012). Teachers, please learn our names!: Racial mi-

- croaggressions and the K-12 classroom. *Race ethnicity and education*, 15(4):441–62. doi:10.1080/13613324.2012.674026
- Kroskrity, P. (1998). Arizona Tewa Kiva speech as a manifestation of a dominant language ideology. In Schieffelin, B., Woolard, K., and Kroskrity, P., editors, *Language ideologies*, pages 103–122. New York: Oxford University Press.
- Kroskrity, P. (2004). Language ideologies. In *A companion to linguistic anthropology*, pages 496–517. Oxford: Blackwell.
- Kubozono, H. (2006). Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua*, 116(7):1140–70. doi:10.1016/j.lingua.2005.06.010
- Labov, W. (1963). The social motivation of sound change. *Word*, 19(3):273–309. doi:10.1080/00437956.1963.11659799
- Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- LaCharité, D. and Paradis, C. (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic inquiry*, 36:223–58. doi:10.1162/0024389053710666
- Lambert, W. E., Hodgson, R. C., Gardner, R. C., and Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The journal of abnormal and social psychology*, 60(1):44–51. doi:10.1037/h0044430
- Laternus, R. (2018). *Perceptual adaptation to non-native speech: The effects of bias, exposure, and input variation*. PhD thesis, New York University.
- Le Page, R. B. and Tabouret-Keller, A. (1985). *Acts of identity: Creole-based approaches to language and identity*. Cambridge: Cambridge University Press.

- Lee, C. (2009, July 3). Barack Obama, a stickler for pronunciation. *Politico*. URL <http://www.politico.com/news/stories/0709/24466.html>
- Lev-Ari, S. and Peperkamp, S. (2014). An experimental study of the role of social factors in language change: The case of loanword adaptations. *Laboratory phonology*, 5(3):379–401. doi:10.1515/lp-2014-0013
- Levay, K., Freese, J., and Druckman, J. (2016). The demographic and political composition of Mechanical Turk samples. *SAGE Open*, 6(1). doi:10.1177/2158244016636433
- Levenson, E. (2014, June 5). A roundup of Obama’s most un-American pronunciations. *The Atlantic*. URL <http://www.theatlantic.com/entertainment/archive/2014/06/a-roundup-of-obamas-most-un-american-pronunciations/372248/>
- Lin, Y. (2008). Variable vowel adaptation in Standard Mandarin loanwords. *Journal of East Asian linguistics*, 17:347–62. doi:10.1007/s10831-008-9031-y
- Lipski, J. (1976). Prejudice and pronunciation. *American speech*, 51:109–18. doi:10.2307/455361
- Lombardi, L. (2003). Second language data and constraints on Manner: Explaining substitutions for the English interdental. *Second language research*, 19(3):225–50. doi:10.1177/026765830301900304
- Los Angeles Times* staff (2016, August 31). Transcript: Donald Trump’s full immigration speech, annotated. *Los Angeles Times*. URL <http://www.latimes.com/politics/la-na-pol-donald-trump-immigration-speech-transcript-20160831-snap-htlstory.html>
- Lovins, J. B. (1975). *Loanwords and the phonological structure of Japanese*. Bloomington: Indiana University Linguistics Club.



- Massey, D. and Lundy, G. (2001). Use of Black English and racial discrimination in urban housing markets: New methods and findings. *Urban affairs review*, 36(4):452–69. doi:10.1177/10780870122184957
- McCarthy, C. (2016, April 27). Trump devises ‘zany’ way to pronounce Tanzania in foreign policy speech. *The Guardian*. URL <http://www.theguardian.com/us-news/2016/apr/27/donald-trump-tanzania-foreign-policy-speech>
- Merry, R. (2016, May 4). Trump vs. Hillary is nationalism vs. globalism, 2016. *The National Interest*. URL <http://nationalinterest.org/feature/trump-vs-hillary-nationalism-vs-globalism-2016-16041>
- Miao, R. (2006). *Loanword adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factors*. PhD thesis, State University of New York, Stony Brook.
- Milroy, L. and Gordon, M. (2003). *Sociolinguistics: Method and interpretation*. Oxford: Blackwell.
- Milroy, L. and Milroy, J. (1992). Social network and social class: Toward an integrated sociolinguistic model. *Language in society*, 21(1):1–26. doi:10.1017/S0047404500015013
- Mindiola, T. and Gutierrez, A. (1988). Chicanos and the legislative process. In García, C., editor, *Latinos and the political system*, pages 349–62. Notre Dame, IN: University of Notre Dame Press.
- Mitchell, B. and Robinson, F. (2011). *A Guide to Old English*. Hoboken, NJ: Wiley, 8 edition.
- Monaghan, P., Shillcock, R., Christiansen, M., and Kirby, S. (2016). How arbitrary is language? *Philosophical transactions of the Royal Society B*, 369:20130299. doi:10.1098/rstb.2013.0299

- Myers-Scotton, C. (1993). Common and uncommon ground: Social and structural factors in codeswitching. *Language in society*, 22(4):475–503. doi:10.1017/S0047404500017449
- Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Nacaskul, K. (1979). A note on English loanwords in Thai. In Thongkum, T., Kullavanijaya, P., Panupong, V., and Tingsabath, M., editors, *Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugénie J.A. Henderson*, pages 151–62. Bangkok: Chulalongkorn University Press.
- Nádasdy, A. (1989). Consonant length in recent borrowings into Hungarian. *Acta linguistica Hungarica*, 39:195–213.
- Nagy, N. (1996). *Language contact and language change in the Faetar speech community*. PhD thesis, University of Pennsylvania.
- Namy, L., Nygaard, L., and Sauerteig, D. (2002). Gender differences in vocal accommodation: The role of perception. *Journal of language and social psychology*, 21(4):422–432. doi:10.1177/026192702237958
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of language and social psychology*, 18(1):62–85. doi:10.1177/0261927X99018001005
- Nosek, B., Smyth, F., Hansen, J., Devos, T., Lindner, N., Ranganath, K., Smith, C., Olson, K., Chugh, D., Greenwald, A., et al. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European review of social psychology*, 18(1):36–88. doi:10.1080/10463280701489053
- Nurse, D. (1985). Dentality, areal features, and phonological change in Northeastern Bantu. *Studies in African linguistics*, 16:243–79.

- Okimoto, T. and Gromet, D. (2016). Differences in sensitivity to deviance partly explain ideological divides in social policy support. *Journal of personality and social psychology*, 111(1):98–117. doi:10.1037/pspp0000080
- Oswald, F., Mitchell, G., Blanton, H., Jaccard, J., and P, T. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of personality and social psychology*, 105(2):171–92. doi:10.1037/a0032734
- Palin, S. (2015, September 6). *State of the Union* [Television broadcast]. Atlanta, GA: CNN. URL <http://www.theguardian.com/us-news/2015/sep/07/sarah-palin-says-immigrants-to-the-us-should-speak-american>
- Pantos, A. and Perkins, A. (2013). Measuring implicit and explicit attitudes toward foreign accented speech. *Journal of language and social psychology*, 32(1):3–20. doi:10.1177/0261927X12463005
- Paradis, C. and LaCharité, D. (1997). Preservation and minimality in loanword adaptation. *Journal of linguistics*, 33(2):379–430. doi:10.1017/S0022226797006786
- Pardo, J. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119(4):2382–2393. doi:10.1121/1.2178720
- Peirce, J. (2016). PsychoPy - Psychophysics Software for Python (manual, release 1.84.2). URL <http://www.psychopy.org/PsychoPyManual.pdf>
- Peperkamp, S., Vendelin, I., and Nakamura, K. (2008). On the perceptual origin of loanword adaptations: Experimental evidence from Japanese. *Phonology*, 25(1):129–164. doi:10.1017/S0952675708001425
- Perniss, P., Thompson, R., and Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in psychology*, 1:227. doi:10.3389/fpsyg.2010.00227

- Petrovic, J. (2005). The conservative restoration and neoliberal defenses of bilingual education. *Language policy*, 4:395–416. doi:10.1007/s10993-005-2888-y
- Pew Research Center staff (2017, April 17). Public dissatisfaction with Washington weighs on the GOP. *Pew Research Center*. URL <http://www.people-press.org/2017/04/17/public-dissatisfaction-with-washington-weighs-on-the-gop/>
- Pew Research Center staff (2018, March 1). The generation gap in American politics. *Pew Research Center*. URL <http://www.people-press.org/2018/03/01/the-generation-gap-in-american-politics/>
- Phillipson, R. (1992). *Linguistic imperialism*. New York: Oxford University Press.
- Phillipson, R. (2009). *Linguistic imperialism continued*. New York: Routledge.
- Piller, I. (2001). Identity constructions in multilingual advertising. *Language in society*, 30(2):153–86.
- Politico (2016). 2016 presidential election results. *Politico*. URL <http://www.politico.com/mapdata-2016/2016-election/results/map/president/>
- Poplack, S. and Dion, N. (2012). Myths and facts about loanword development. *Language variation and change*, 24(3):279–315. doi:10.1017/S095439451200018X
- Poplack, S. and Sankoff, D. (1984). Borrowing: The synchrony of integration. *Linguistics*, 22(1):99–136. doi:10.1515/ling.1984.22.1.99
- Poplack, S., Sankoff, D., and Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Journal of linguistics*, 26(1):47–104. doi:10.1515/ling.1988.26.1.47
- Pratto, F., Sidanius, J., L, S., and B, M. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of personality and social psychology*, 67:741–763. doi:10.1037/0022-3514.67.4.741

- Preston, D. (1989). *Perceptual dialectology: Nonlinguists' views of areal linguistics*. Providence, RI: Foris Publications.
- Preston, D. (1996). Whaddayaknow?: The modes of folk linguistic awareness. *Language awareness*, 5(1):40–74. doi:10.1080/09658416.1996.9959890
- Preston, D. (1999). A language attitude analysis of regional US speech: Is Northern US English not friendly enough? *Cuadernos de filología inglesa*, 8:129–46.
- Project Implicit (2016). Project Implicit. URL <https://implicit.harvard.edu/implicit/>
- Purnell, T., Idsardi, W., and Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of language and social psychology*, 18(1):10–30. doi:10.1177/0261927X99018001002
- Radeau, M., Morais, J., and Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of experimental psychology*, 21(6):1297–1311. doi:10.1037/0096-1523.21.6.1297
- RCoreTeam (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ren, H. (1986). On the acoustic structure of diphthongal syllables. In *UCLA Working Papers in Phonetics*, volume 65. UCLA.
- Ricento, T. (2005). Problems with the "language as resource" discourse in the promotion of heritage languages in the U.S.A. *Journal of sociolinguistics*, 9:348–68. doi:10.1111/j.1360-6441.2005.00296.x
- Rickford, J. (1986). The need for new approaches to social class analysis in sociolinguistics. *Language and communication*, 6(3):215–21.

- Rickford, J. and McNair-Knox, F. (1994). Addressee- and topic-influenced style shift: A quantitative sociolinguistic study. In Biber, D. and Finegan, E., editors, *Sociolinguistic perspectives on register*, pages 235–276. Oxford: Oxford University Press.
- Rose, Y. and Demuth, K. (2006). Vowel epenthesis in loanword adaptation: Representational and phonetic considerations. *Lingua*, 116(7):1112–39. doi:10.1016/j.lingua.2005.06.011
- Rothstein, R. (2001). Bilingual education: The controversy. In *The Jossey-Bass reader on school reform*, chapter 15, pages 218–232. San Francisco: Jossey-Bass.
- Ruiz, R. (1984). Orientations in language planning. *NABE journal*, 8(2):15–34. doi:10.1080/08855072.1984.10668464
- Saciuk, B. (1969). The stratal division of the lexicon. *Linguistics*, 1:464–532. doi:10.1080/08351816909389128
- Safire, W. (2008). *Safire's political dictionary*. Oxford: Oxford University Press.
- San Giacomo, M. and Peperkamp, S. (2008). Presencia del español en náhuatl: Estudio sociolingüístico de la adaptación de préstamos. In Westmoreland, M. and Thomas, J., editors, *Selected proceedings of the Fourth Workshop on Spanish Sociolinguistics*, pages 149–156. Somerville, MA: Cascadilla Press.
- Sankoff, G. (2002). Linguistic outcomes of language contact. In Chambers, J., Trudgill, P., and Schilling-Estes, N., editors, *The handbook of language variation and change*, pages 729–63. Oxford: Blackwell.
- Santoro, W. (1999). Conventional politics takes center stage: The Latino struggle against English-only laws. *Social Forces*, 77(3):887–909. doi:10.1093/sf/77.3.887
- Schilling-Estes, N. (1998). Investigating ‘self-conscious’ speech: The performance register in Ocracoake English. *Language in society*, 27:53–83. doi:10.1017/S0047404500019722

- Schilling-Estes, N. (2008). Stylistic variation and the sociolinguistic interview: A reconsideration. In Monroy, R. and Sánchez, A., editors, *25 años de lingüística aplicada en España: Hitos y retos*, pages 971–986. Murcia: Servicio de Publicaciones de la Universidad de Murcia.
- Schmidt, R. (1998). The politics of language in Canada and the United States: Explaining the differences. In Ricento, T. and Burnaby, B., editors, *Language and politics in the United States and Canada: Myths and realities*, pages 37–70. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Shinohara, S. (1997). *Analyse phonologique de l'adaptation japonaise de mots étrangers*. PhD thesis, Université Sorbonne-Nouvelle.
- Shinohara, S. (2004). Emergence of Universal Grammar in foreign word adaptations. In Kager, R., Pater, J., and Zonneveld, W., editors, *Constraints in phonological acquisition*, pages 292–320. Cambridge: Cambridge University Press.
- Shinohara, S. (2006). Perceptual effects in final cluster reduction patterns. *Lingua*, 116:1046–78. doi:10.1016/j.lingua.2005.06.008
- Sidanius, J. and Pratto, F. (1999). *Social dominance*. Cambridge: Cambridge University Press.
- Silva, D., Peters, S., Duhaish, F. B., Kim, S.-H., Koo, Y., Marji, L., and Park, J. (2011). Variation in the *Iraq* vowels outside the public forum: The indexing of political persuasion reconsidered. *American speech*, 86(2):179–191. doi:10.1215/00031283-1337001
- Silverman, D. (1992). Multiple scansion in loanword phonology: Evidence from Cantonese. *Phonology*, 9(2):298–328. doi:10.1017/S0952675700001627
- Silverstein, M. (1979). Language structure and linguistic ideology. In Clyne, P., Hanks,

- W., and Hofbauer, C., editors, *The elements: A parasession on linguistic units and levels*, pages 193–247. Chicago: Chicago Linguistic Society.
- Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language and communication*, 23(3):193–229. doi:10.1016/S0271-5309(03)00013-2
- Singh, F. (1985). Prosodic adaptation in interphonology. *Lingua*, 67:269–282. doi:10.1016/0024-3841(85)90001-4
- Smith, B. (2009a). Dental fricatives and stops in Germanic: Deriving diachronic processes from synchronic variation. In Dufresne, M., Dupuis, F., and Vocaj, E., editors, *Historical linguistics 2007: Selected papers from the 18th International Conference on Historical Linguistics, Montreal, 6–11 August 2007*, pages 19–36. Amsterdam: John Benjamins.
- Smith, J. L. (2006). Loan phonology is not all perception: Evidence from Japanese loan doublets. In Vance, T. J. and Jones, K. A., editors, *Japanese/Korean Linguistics 14*, pages 63–74.
- Smith, J. L. (2009b). Source similarity in loanword adaptation: Correspondence Theory and the posited source-language representation. In Parker, S., editor, *Phonological Argumentation: Essays on Evidence and Motivation*, pages 155–177. London: Equinox.
- Squires, G. and Chadwick, J. (2006). Linguistic profiling: A continuing tradition of discrimination in the home insurance industry? *Urban affairs review*, 41(3):400–15. doi:10.1177/1078087405281064
- Stafford, T. and Scaife, R. (2015). Open IAT, 2015-09-03 version. *Open Science Framework*. URL <https://osf.io/i8yj5/>
- Steinbergs, A. (1985). The role of MSC's in Oshikwanyama loan phonology. *Studies in African linguistics*, 16:89–101.



- Steriade, D. (2009). The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In Hanson, K. and Inkelas, S., editors, *The nature of the word: Studies in honor of Paul Kiparsky*, pages 151–180. Cambridge, MA: MIT Press.
- Sumner, M. and Samuel, A. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of memory and language*, 60:487–501. doi:10.1016/j.jml.2009.01.001
- Tatalovich, R. (1995). *Nativism reborn?* Lexington, KY: University Press of Kentucky.
- The Economist* staff (2009, June 4). Barack Hussein’s Arabic. *The Economist*. URL <https://www.economist.com/democracy-in-america/2009/06/04/barack-husseins-arabic>
- Thomason, S. G. (2001). *Language contact*. Edinburgh: Edinburgh University Press. doi:10.1016/B0-08-043076-7/03032-1
- Thomason, S. G. (2007). Language contact and deliberate change. *Journal of Language Contact*, 1(1):41–62.
- Thomason, S. G. and Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley, CA: University of California Press.
- Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in society*, 1:179–195. doi:10.1017/S0047404500000488
- Uffmann, C. (2006). Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua*, 116:1079–1111. doi:10.1016/j.lingua.2005.06.009
- Ulrich, C. H. (1997). Loanword adaptation in Lama: Testing the TCRS model. *The Canadian journal of linguistics*, 42(4):415–463. doi:10.1017/S0008413100017163
- Underhill, R. (1988). *Like* is, like, focus. *American speech*, 63(3):234–46. doi:10.2307/454820

- Valdés, G. (1997). Dual language immersion programs: A cautionary note concerning the education of language-minority students. *Harvard educational review*, 67:391–429. doi:10.17763/haer.67.3.n5q175qp86120948
- van Coetsem, F. (1988). *Loan phonology and the two transfer types in language contact*. Berlin: De Gruyter.
- van Coetsem, F. (2000). *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- van Oostendorp, M. (1997). Style levels in conflict resolution. In *Amsterdam studies in the theory and history of linguistic science series*, pages 207–230. Amsterdam: John Benjamins.
- Vendelin, I. and Peperkamp, S. (2006). The influence of orthography on loanword adaptations. *Lingua*, 116(7):996–1007. doi:10.1016/j.lingua.2005.07.005
- Weatherholtz, K., Campbell-Kibler, K., and Jaeger, T. F. (2014). Socially-mediated syntactic alignment. *Language variation and change*, 26(03):387–420. doi:10.1017/S0954394514000155
- Weinberger, S. (1994). Functional and phonetic constraints on second language acquisition. In Yavas, M., editor, *First and second language phonology*, pages 283–302. San Diego: Singular.
- Weinberger, S. (1997). Minimal segments in second language phonology. In James, A. and Leather, J., editors, *Second-language speech: Structure and process*, pages 263–311. Berlin: De Gruyter.
- Weinreich, U. (1968). *Languages in contact: Findings and problems*. The Hague: Mouton. doi:10.1515/9783110802177
- Weinreich, U., Labov, W., and Herzog, M. (1968). Empirical foundations for a theory

- of language change. In Lehmann, W. and Malkiel, Y., editors, *Directions for historical linguistics: A symposium*, pages 95–195. Austin: University of Texas Press.
- Westfall, J., van Boven, L., Chambers, J., and Judd, C. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspectives on psychological science*, 10(2):145–158. doi:10.1177/1745691615569849
- Williams, A., Garrett, P., and Coupland, N. (1999). Dialect recognition. In Preston, D., editor, *Handbook of perceptual dialectology*, pages 345–58. Amsterdam: John Benjamins.
- Winter, B. (2014). A very basic tutorial for performing linear mixed effects analysis (Tutorial 2). *arXiv*, 1308.5499. URL <http://arxiv.org/pdf/1308.5499.pdf>
- Woolard, K. (1985). Language variation and cultural hegemony: Toward an integration of sociolinguistic and social theory. *American ethnologist*, 12(4):738–48. doi:10.1525/ae.1985.12.4.02a00090
- Woolard, K. (1992). Language ideology: Issues and approaches. *Pragmatics*, 2(3):235–249.
- Yi, H., Phelps, J., Smiljanic, R., and Chandrasekaran, B. (2013). Reduced efficiency of audiovisual integration for nonnative speech. *The journal of the Acoustical Society of America*, 134(5):387–393. doi:10.1121/1.4822320
- Yip, M. (1993). Cantonese loanword phonology and Optimality Theory. *Journal of East Asian linguistics*, 2:261–91. doi:10.1007/BF01739135
- Yip, M. (2006). The symbiosis between perception and grammar in loanword phonology. *Lingua*, 116:950–975. doi:10.1016/j.lingua.2005.05.007
- Yu, A., Abrego-Collier, C., and Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and “autistic” traits. *PLoS ONE*, 8(9):e74746. doi:10.1371/journal.pone.0074746

- Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American speech*, 85(3):315–337. doi:10.1215/00031283-2010-018
- Zahn, C. and Hopper, R. (1985). Measuring language attitudes: The speech evaluation instrument. *Journal of language and social psychology*, 4(2):113–23. doi:10.1177/0261927X8500400203
- Zhang, Q. (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in society*, 34(3):431–66. doi:10.1017/S0047404505050153