Appendix: Supplementary materials

A1. Procedures

Upon providing consent to take part in the experiment, each participant was led into a small room that contained a computer, a set of headphones, and a microphone. The experimenter gave the following oral instructions to each participant:

> In this experiment, you're going to use the headphones to hear words from a language. In the first part of the experiment, there is a short section where you will learn a little bit about how the language works. In the second part of the experiment, you'll hear a word, repeat it out loud, and then you can press any key on the keyboard to hear the next one, and so on. The microphone is going to record during that part of the experiment, but those recordings won't be used for anything outside of this study. Then there is a final section where you'll be tested on what you've learned, and for that part you'll switch to using the button box. You're going to hear two options for the right answer. If you think the first one is right, press '1,' and if you think the second one is right, press '2.' We ask that you use one finger from each hand when selecting your answer.

Specific instructions were repeated with on-screen text prior to each phase of the experiment. After completing the study, subjects completed a language-background questionnaire used for the purpose of determining whether or not their results were eligible for inclusion in the statistical analyses.

A2. Stimuli

Table A1: Training stimuli for stems with no liquids.

Stem	Past (-li)	Future (-ru)	Speaker	Block
tikemu	tikemuli	tikemuru	1	1
kibupi	kibupili	kibupiru	1	1
pupugu	pupuguli	pupuguru	1	1
gonuni	gonunili	gonuniru	1	1
bipobe	bipobeli	bipoberu	1	1
tepobi	tepobili	tepobiru	1	1
tomeku	tomekuli	tomekuru	1	2
pibogo	pibogoli	pibogoru	1	2
nekine	nekineli	nekineru	1	2
mutumu	mutumuli	mutumuru	1	2
dubope	dubopeli	duboperu	1	2
degiti	degitili	degitiru	1	2
kukedo	kukedoli	kukedoru	1	1
nomene	nomeneli	nomeneru	1	1
gegebi	gegebili	gegebiru	1	1
butopi	butopili	butopiru	1	1
dodigo	dodigoli	dodigoru	1	1
nimimo	nimimoli	nimimoru	1	1
pededu	pededuli	pededuru	1	2
minoko	minokoli	minokoru	1	2
gutudo	gutudoli	gutudoru	1	2
mogiku	mogikuli	mogikuru	1	2
konute	konuteli	konuteru	1	2
bedite	beditеli	beditеru	1	2
podoge	podogeli	podogeru	2	1
gibipe	gibipeli	gibiperu	2	1
topidu	topiduli	topiduru	2	1
potetu	potetuli	poteturu	2	1
bumumo	bumumoli	bumumoru	2	1
dimumi	dimumili	dimumiru	2	1
botini	botinili	botiniru	2	2
gipebu	gipebuli	gipeburu	2	2
denenu	denenuli	denenuru	2	2
nupidi	nupidili	nupidiru	2	2
bokuno	bokunoli	bokunoru	2	2
nonegu	noneguli	noneguru	2	2
kemoti	kemotili	kemotiru	2	1
digupo	digupoli	diguporu	2	1
pubigi	pubigili	pubigiru	2	1

Continued on next page

(*Continued from previous page*)

Table A1 – Training stimuli (no liquids)

Stem	Past (-li)	Future (-ru)	Speaker	Block
medoto	medotoli	medotoru	2	1
tuniki	tunikili	tunikiru	2	1
kekoke	kekokeli	kekokeru	2	1
megobe	megobeli	megoberu	2	2
gebepu	gebepuli	gebepuru	2	2
migode	migodeli	migoderu	2	2
nukuko	nukukoli	nukukoru	2	2
kidubo	kiduboli	kiduboru	2	2
tuteme	tutemeli	tutemeru	2	2
pemoti	pemotili	pemotiru	3	1
tetobu	tetobuli	tetoburu	3	1
tetige	tetigeli	tetigeru	3	1
begiku	begikuli	begikuru	3	1
pipeto	pipetoli	pipetoru	3	1
gokine	gokineli	gokineru	3	1
monube	monubeli	monuberu	3	2
dobemu	dobemuli	dobemuru	3	2
nituki	nitukili	nitukiru	3	2
mupudu	mupuduli	mupuduru	3	2
pibupi	pibupili	pibupiru	3	2
doduno	dodunoli	dodunoru	3	2
kugumi	kugumili	kugumiru	3	1
kugeko	kugekoli	kugekoru	3	1
nemide	nemideli	nemideru	3	1
mubope	mubopeli	muboperu	3	1
bikote	bikoteli	bikoteru	3	1
tidido	tididoli	tididoru	3	1
gomepu	gomepuli	gomepuru	3	2
duponi	duponili	duponiru	3	2
kodegu	kodeguli	kodeguru	3	2
nenobi	nenobili	nenobiru	3	2
gukego	gukegoli	gukegoru	3	2
binimo	binimoli	binimoru	3	2
mipede	mipedeli	mipederu	4	1
pegono	pegonoli	pegonoru	4	1
kikuge	kikugeli	kikugeru	4	1
dibigi	dibigili	dibigiru	4	1
bukoke	bukokeli	bukokeru	4	1
ninipu	ninipuli	ninipuru	4	1
bemubo	bemuboli	bemuboru	4	2

Continued on next page

(*Continued from previous page*)

Table A1 – Training stimuli (no liquids)

Stem	Past (-li)	Future (-ru)	Speaker	Block
kedetu	kedetuli	kedeturu	4	2
bobeki	bobekili	bobekiru	4	2
gepunu	gepunuli	gepunuru	4	2
podupo	podupoli	poduporu	4	2
metopo	metopoli	metoporu	4	2
nodobu	nodobuli	nodoburu	4	1
titeme	titemeli	titemeru	4	1
kotike	kotikeli	kotikeru	4	1
nubutu	nubutuli	nubuturu	4	1
momime	momimeli	momimeru	4	1
ginoto	ginotoli	ginotoru	4	1
tunenu	tunenuli	tunenuru	4	2
gukidi	gukidili	gukidiru	4	2
pumogi	pumogili	pumogiru	4	2
togudi	togudili	togudiru	4	2
degemi	degemili	degemiru	4	2
dupibo	dupiboli	dupiboru	4	2

Table A2: Training stimuli for 'Short-range' stems

Stem	Harm (-li)	Harm (-ru)	Diss (-li)	Diss (-ru)	Speaker	Liquid	Block
pupoli	pupolili	puporiru	puporili	pupoliru	1	L	1
depile	depileli	depireru	depireli	depileru	1	L	1
nobilo	nobiloli	nobiroru	nobiroli	nobiloru	1	L	1
kinelu	kineluli	kineruru	kineruli	kineluru	1	L	1
bubeli	bubelili	buberiru	buberili	bubeliru	1	L	1
gepele	gepeleli	gepereru	gepereli	gepeleru	1	L	1
petili	petilili	petiriru	petirili	petiliru	1	L	2
ninole	ninoleli	ninoreru	ninoreli	ninoleru	1	L	2
tonulo	tonuloli	tonuroru	tonuroli	tonuloru	1	L	2
kogilu	kogiluli	kogiruru	kogiruli	kogiluru	1	L	2
mimelo	mimeloli	mimeroru	mimeroli	mimeloru	1	L	2
gukolu	gukoluli	gukoruru	gukoruli	gukoluru	1	L	2
giberi	gibelili	giberiru	giberili	gibeliru	1	R	1
kemure	kemuleli	kemureru	kemureli	kemuleru	1	R	1
pomoro	pomololi	pomororu	pomoroli	pomoloru	1	R	1
boguru	bogululi	bogururu	boguruli	boguluru	1	R	1
nunuro	nunuloli	nunuroru	nunuroli	nunuloru	1	R	1
digiru	digiluli	digiruru	digiruli	digiluru	1	R	1
tuguri	tugulili	tuguriru	tugurili	tuguliru	1	R	2
dupire	dupileli	dupireru	dupireli	dupileru	1	R	2
bebiro	bebiloli	bebiroru	bebiroli	bebiloru	1	R	2
tineru	tineluli	tineruru	tineruli	tineluru	1	R	2
metiri	metilili	metiriru	metirili	metiliru	1	R	2
mopere	mopeleli	mopereru	mopereli	mopeleru	1	R	2
duguli	dugulili	duguriru	dugurili	duguliru	2	L	1
gumule	gumuleli	gumureru	gumureli	gumuleru	2	L	1
bimolo	bimololi	bimororu	bimoroli	bimoloru	2	L	1
negulu	negululi	negururu	neguruli	neguluru	2	L	1
piboli	pibolili	piboriru	piborili	piboliru	2	L	1
kokelu	kokeluli	kokeruru	kokeruli	kokeluru	2	L	1
todeli	todelili	toderiru	toderili	todeliru	2	L	2
modile	modileli	modireru	modireli	modileru	2	L	2
bunilo	buniloli	buniroru	buniroli	buniloru	2	L	2
tetilu	tetiluli	tetiruru	tetiruli	tetiluru	2	L	2
detele	deteleli	detereru	detereli	deteleru	2	L	2
mibulo	mibuloli	miburoru	miburoli	mibuloru	2	L	2
kupori	kupolili	kuporiru	kuporili	kupoliru	2	R	1
gonore	gonoleli	gonoreru	gonoreli	gonoleru	2	R	1
nimero	nimeloli	nimeroru	nimeroli	nimeloru	2	R	1
pekoru	pekoluli	pekoruru	pekoruli	pekoluru	2	R	1
beture	betuleli	betureru	betureli	betuleru	2	R	1

Continued on next page

Table A2 – Training stimuli (Short-range) (*Continued from previous page*)

Stem	Harm (-li)	Harm (-ru)	Diss (-li)	Diss (-ru)	Speaker	Liquid	Block
dogoro	dogololi	dogororu	dogoroli	dogoloru	2	R	1
kideri	kidelili	kideriru	kiderili	kideliru	2	R	2
nodire	nodileli	nodireru	nodireli	nodileru	2	R	2
muburo	mubuloli	muburoru	muburoli	mubuloru	2	R	2
gemuru	gemululi	gemururu	gemuruli	gemuluru	2	R	2
pukuri	pukulili	pukuriru	pukurili	pukuliru	2	R	2
titiru	titiluli	titiruru	titiruli	titiluru	2	R	2
pogili	pogilili	pogiriru	pogirili	pogiliru	3	L	1
dikole	dikoleli	dikoreru	dikoreli	dikoleru	3	L	1
ginelo	gineloli	gineroru	gineroli	gineloru	3	L	1
numolu	numoluli	numororu	numoruli	numoluru	3	L	1
bogolo	bogololi	bogororu	bogoroli	bogoloru	3	L	1
kemulu	kemululi	kemururu	kemuruli	kemuluru	3	L	1
pokuli	pokulili	pokuriru	pokurili	pokuliru	3	L	2
gitule	gituleli	gitureru	gitureli	gituleru	3	L	2
tubilo	tubiloli	tubiroru	tubiroli	tubiloru	3	L	2
mepelu	mepeluli	meperuru	meperuli	mepeluru	3	L	2
tutoli	tutolili	tutoriru	tutorili	tutoliru	3	L	2
memile	memileli	memireru	memireli	memileru	3	L	2
gegiri	gegilili	gegiriru	gegirili	gegiliru	3	R	1
notere	noteleli	notereru	notereli	noteleru	3	R	1
puniro	puniloli	puniroru	puniroli	puniloru	3	R	1
dekeru	dekeluli	dekeruru	dekeruli	dekeluru	3	R	1
kibori	kibolili	kiboriru	kiborili	kiboliru	3	R	1
bukore	bukoleli	bukoreru	bukoreli	bukoleru	3	R	1
digeri	digelili	digeriru	digerili	digeliru	3	R	2
topure	topuleli	topureru	topureli	topuleru	3	R	2
bonero	boneloli	boneroru	boneroli	boneloru	3	R	2
mimoru	mimoluli	mimororu	mimoruli	mimoluru	3	R	2
nutoro	nutololi	nutororu	nutoroli	nutoloru	3	R	2
keduru	kedululi	kedururu	keduruli	keduluru	3	R	2
begeli	begelili	begeriru	begerili	begeliru	4	L	1
gokele	gokeleli	gokereru	gokereli	gokeleru	4	L	1
kubulo	kubuloli	kuburoru	kuburoli	kubuloru	4	L	1
nidilu	nidiluli	nidiruru	nidiruli	nidiluru	4	L	1
pedole	pedoleli	pedoreru	pedoreli	pedoleru	4	L	1
domelo	domeloli	domeroru	domeroli	domeloru	4	L	1
bikili	bikilili	bikiriru	bikirili	bikiliru	4	L	2
tipule	tipuleli	tipureru	tipureli	tipuleru	4	L	2
dotolo	dotololi	dotororu	dotoroli	dotoloru	4	L	2
kupolu	kupoluli	kuporuru	kuporuli	kupoluru	4	L	2
muduli	mudulili	muduriru	mudurili	muduliru	4	L	2

Continued on next page

Table A2 – Training stimuli (Short-range) (*Continued from previous page*)

Stem	Harm (-li)	Harm (-ru)	Diss (-li)	Diss (-ru)	Speaker	Liquid	Block
nedulu	nedululi	nedururu	neduruli	neduluru	4	L	2
biduri	bidulili	biduriru	bidurili	biduliru	4	R	1
komire	komileli	komireru	komireli	komileru	4	R	1
guburo	gubuloli	guburoru	guburoli	gubuloru	4	R	1
dudiru	dudiluli	dudiruru	dudiruli	dudiluru	4	R	1
nekiri	nekilili	nekiriru	nekirili	nekiliru	4	R	1
piperu	pipeluli	piperuru	piperuli	pipeluru	4	R	1
tetori	tetolili	tetoriru	tetorili	tetoliru	4	R	2
pikere	pikeleli	pikereru	pikereli	pikeleru	4	R	2
gobiro	gobiloli	gobiroru	gobiroli	gobiloru	4	R	2
moporu	mopoluli	moporuru	moporuli	mopoluru	4	R	2
tedore	tedoleli	tedoreru	tedoreli	tedoleru	4	R	2
mumero	mumeloli	mumeroru	mumeroli	mumeloru	4	R	2

Table A3: Training stimuli for 'Medium-range' stems

Stem	Harm (-li)	Harm (-ru)	Diss (-li)	Diss (-ru)	Speaker	Liquid	Block
polipu	polipuli	poripuru	poripuli	polipuru	1	L	1
pilede	piledeli	pirederu	piredeli	pilederu	1	L	1
bilono	bilonoli	bironoru	bironoli	bilonoru	1	L	1
neluki	nelukili	nerukiru	nerukili	nelukiru	1	L	1
belibu	belibuli	beriburu	beribuli	beliburu	1	L	1
pelege	pelegeli	peregeru	peregeli	pelegeru	1	L	1
tilipe	tilipeli	tiriperu	tiripeli	tiliperu	1	L	2
noleni	nolenili	noreniru	norenili	noleniru	1	L	2
nuloto	nulotoli	nurotoru	nurotoli	nulotoru	1	L	2
giluko	gilukoli	girukoru	girukoli	gilukoru	1	L	2
melomi	melomili	meromiru	meromili	melomiru	1	L	2
kolugu	koluguli	koruguru	koruguli	koluguru	1	L	2
berigi	beligili	berigiru	berigili	beligiru	1	R	1
mureke	mulekeli	murekeru	murekeli	mulekeru	1	R	1
moropo	molopoli	moroporu	moropoli	moloporu	1	R	1
gurubo	guluboli	guruboru	guruboli	guluboru	1	R	1
nuronu	nulonuli	nuronuru	nuronuli	nulonuru	1	R	1
girudi	giludili	girudiru	girudili	giludiru	1	R	1
guritu	gulituli	gurituru	gurituli	gulituru	1	R	2
piredu	pileduli	pireduru	pireduli	pileduru	1	R	2
birobe	bilobeli	biroberu	birobeli	biloberu	1	R	2
neruti	nelutili	nerutiru	nerutili	nelutiru	1	R	2
tirime	tilimeli	tirimeru	tirimeli	tilimeru	1	R	2
peremo	pelemoli	peremoru	peremoli	pelemoru	1	R	2
gulidu	guliduli	guriduru	guriduli	guliduru	2	L	1
mulegu	muleguli	mureguru	mureguli	muleguru	2	L	1
molobi	molobili	morobiru	morobili	molobiru	2	L	1
gulune	guluneli	guruneru	guruneli	guluneru	2	L	1
bolipi	bolipili	boripiru	boripili	bolipiru	2	L	1
keluko	kelukoli	kerukoru	kerukoli	kelukoru	2	L	1
delito	delitoli	deritoru	deritoli	delitoru	2	L	2
dilemo	dilemoli	diremoru	diremoli	dilemoru	2	L	2
nilobu	nilobuli	niroburu	nirobuli	niloburu	2	L	2
tilute	tiluteli	tiruteru	tiruteli	tiluteru	2	L	2
telede	teledeli	terederu	teredeli	telederu	2	L	2
bulomi	bulomili	buromiru	buromili	bulomiru	2	L	2
poriku	polikuli	porikuru	porikuli	polikuru	2	R	1
norego	nolegoli	noregoru	noregoli	nolegoru	2	R	1
meroni	melonili	meroniru	meronili	meloniru	2	R	1
korupe	kolupeli	koruperu	korupeli	koluperu	2	R	1
turebe	tulebeli	tureberu	turebeli	tuleberu	2	R	1

Continued on next page

Table A3 – Training stimuli (Medium-range) (*Continued from previous page*)

Stem	Harm (-li)	Harm (-ru)	Diss (-li)	Diss (-ru)	Speaker	Liquid	Block
gorodo	golodoli	gorodoru	gorodoli	golodoru	2	R	1
deriki	delikili	derikiru	derikili	delikiru	2	R	2
direno	dilenoli	direnoru	direnoli	dilenoru	2	R	2
buromu	bulomuli	buromuru	buromuli	bulomuru	2	R	2
muruge	mulugeli	murugeru	murugeli	mulugeru	2	R	2
kuripu	kulipuli	kuripuru	kuripuli	kulipuru	2	R	2
tiruti	tilutili	tirutiru	tirutili	tilutiru	2	R	2
gilipo	gilipoli	giriporu	giripoli	giliporu	3	L	1
koledi	koledili	korediru	koredili	kolediru	3	L	1
nelogi	nelogili	nerogiru	nerogili	nelogiru	3	L	1
molunu	molunuli	morunuru	morunuli	molunuru	3	L	1
golobo	goloboli	goroboru	goroboli	goloboru	3	L	1
muluke	mulukeli	murukeru	murukeli	mulukeru	3	L	1
kulipo	kulipoli	kuriporu	kuripoli	kuliporu	3	L	2
tulegi	tulegili	turegiru	turegili	tulegiru	3	L	2
bilotu	bilotuli	biroturu	birotuli	biloturu	3	L	2
pelume	pelumeli	perumeru	perumeli	pelumeru	3	L	2
tolitu	tolituli	torituru	torituli	tolituru	3	L	2
mileme	milemeli	miremeru	miremeli	milemeru	3	L	2
girige	giligeli	girigeru	girigeli	giligeru	3	R	1
tereno	telenoli	terenoru	terenoli	telenoru	3	R	1
niropu	nilopuli	niropuru	niropuli	nilopuru	3	R	1
kerude	keludeli	keruderu	kerudeli	keluderu	3	R	1
boriki	bolikili	borikiru	borikili	bolikiru	3	R	1
korebu	kolebuli	koreburu	korebuli	koleburu	3	R	1
geridi	gelidili	geridiru	geridili	gelidiru	3	R	2
pureto	puletoli	puretoru	puretoli	puletoru	3	R	2
nerobo	neloboli	neroboru	neroboli	neloboru	3	R	2
morumi	molumili	morumiru	morumili	molumiru	3	R	2
toronu	tolonuli	toronuru	toronuli	tolonuru	3	R	2
duruke	dulukeli	durukeru	durukeli	dulukeru	3	R	2
gelibe	gelibeli	geriberu	geribeli	geliberu	4	L	1
kelego	kelegoli	keregoru	keregoli	kelegoru	4	L	1
buloku	bulokuli	burokuru	burokuli	bulokuru	4	L	1
diluni	dilunili	diruniru	dirunili	diluniru	4	L	1
dolepe	dolepeli	doreperu	dorepeli	doleperu	4	L	1
melodo	melodoli	merodoru	merodoli	melodoru	4	L	1
kilibi	kilibili	kiribiru	kiribili	kilibiru	4	L	2
puleti	puletili	puretiru	puretili	puletiru	4	L	2
tolodo	tolodoli	torodoru	torodoli	tolodoru	4	L	2
poluku	polukuli	porukuru	porukuli	polukuru	4	L	2
dulimu	dulimuli	durimuru	durimuli	dulimuru	4	L	2

Continued on next page

Table A3 – Training stimuli (Medium-range) (*Continued from previous page*)

Stem	Harm (-li)	Harm (-ru)	Diss (-li)	Diss (-ru)	Speaker	Liquid	Block
dulune	duluneli	duruneru	duruneli	duluneru	4	L	2
duribi	dulibili	duribiru	duribili	dulibiru	4	R	1
mireko	milekoli	mirekoru	mirekoli	milekoru	4	R	1
burogu	buloguli	buroguru	buroguli	bulоguru	4	R	1
dirudu	diluduli	diruduru	diruduli	diluduru	4	R	1
kirine	kilineli	kirineru	kirineli	kilineru	4	R	1
perupi	pelupili	perupiru	perupili	pelupiru	4	R	1
torite	toliteli	toriteru	toriteli	toliteru	4	R	2
kerepi	kelepili	kerepiru	kerepili	kelepiru	4	R	2
birogo	bilogoli	birogoru	birogoli	bilogoru	4	R	2
porumo	polumoli	porumoru	porumoli	polumoru	4	R	2
dorete	doleteli	doreteru	doreteli	doleteru	4	R	2
meromu	melomuli	meromuru	meromuli	melomuru	4	R	2

Table A4: Stimuli used in the testing phase of all experiments.

Distance	Stem	Option 1	Option 2	Speaker	Liquid	Suffix	Block	Faith	First
Short	dotile	dotileli	dotireli	1	L	li	1	H	H
Short	tipoli	tipolili	tiporili	1	L	li	2	H	H
Short	pidole	pidoleru	pidoreru	1	L	ru	1	D	D
Short	nonolu	nonoluru	nonoruru	1	L	ru	2	D	D
Short	gegori	gegolili	gegorili	1	R	li	1	D	H
Short	kidure	kiduleli	kidureli	1	R	li	2	D	H
Short	mipuru	mipuluru	mipururu	1	R	ru	1	H	D
Short	ditore	ditoleru	ditoreru	1	R	ru	2	H	D
Short	bibolo	biboloili	biboroli	2	L	li	1	H	H
Short	pudele	pudeleli	pudereli	2	L	li	2	H	H
Short	tepilo	tepiloru	tepiroru	2	L	ru	1	D	D
Short	gigili	gigiliru	gigiriru	2	L	ru	2	D	D
Short	dutere	duteleli	dutereli	2	R	li	1	D	H
Short	popero	popeloli	poperoli	2	R	li	2	D	H
Short	pemeri	pemeliru	pemeriru	2	R	ru	1	H	D
Short	goniro	goniloru	goniroru	2	R	ru	2	H	D
Short	guneli	gunelili	gunerili	3	L	li	1	H	H
Short	momilu	momiluli	momiruli	3	L	li	2	H	H
Short	mukelu	mukeluru	mukeruru	3	L	ru	1	D	D
Short	detule	detuleru	detureru	3	L	ru	2	D	D
Short	nibiru	nibiluli	nibiruli	3	R	li	1	D	H
Short	memoru	memoluli	memoruli	3	R	li	2	D	H
Short	kudire	kudileru	kudireru	3	R	ru	1	H	D
Short	nuburi	nubuliru	nuburiru	3	R	ru	2	H	D
Short	negulu	negululi	neguruli	4	L	li	1	H	H
Short	kekulo	kekuloli	kekuroli	4	L	li	2	H	H
Short	komuli	komuliru	komuriru	4	L	ru	1	D	D
Short	bubelo	bubeloru	buberoru	4	L	ru	2	D	D
Short	bonuro	bonuloli	bonuroli	4	R	li	1	D	H
Short	tukiri	tukilili	tukirili	4	R	li	2	D	H
Short	tokoro	tokoloru	tokororu	4	R	ru	1	H	D
Short	begeru	begeluru	begeruru	4	R	ru	2	H	D
Medium	beliki	belikili	berikili	1	L	li	1	H	H
Medium	dilopo	dilopoli	diropoli	1	L	li	2	H	H
Medium	pilepe	pileperu	pireperu	1	L	ru	1	D	D
Medium	muluto	mulutoru	murutoru	1	L	ru	2	D	D
Medium	burike	bulikeli	burikeli	1	R	li	1	D	H
Medium	dorupi	dolupili	dorupili	1	R	li	2	D	H
Medium	porodi	polodiru	porodiru	1	R	ru	1	H	D
Medium	mirete	mileteru	mireteru	1	R	ru	2	H	D
Medium	molutu	molutuli	morutuli	2	L	li	1	H	H

Continued on next page

Table A4 – Testing stimuli (*Continued from previous page*)

Distance	Stem	Option 1	Option 2	Speaker	Liquid	Suffix	Block	Faith	First
Medium	pelemi	pelemili	peremili	2	L	li	2	H	H
Medium	nulimu	nulimuru	nurimuru	2	L	ru	1	D	D
Medium	tolone	toloneru	toroneru	2	L	ru	2	D	D
Medium	merenu	melenuli	merenuli	2	R	li	1	D	H
Medium	puremo	pulemoli	puremoli	2	R	li	2	D	H
Medium	girupu	gilupuru	girupuru	2	R	ru	1	H	D
Medium	teriko	telikoru	terikoru	2	R	ru	2	H	D
Medium	kilono	kilonoli	kironoli	3	L	li	1	H	H
Medium	gulibe	gulibeli	guribeli	3	L	li	2	H	H
Medium	golobi	golobiru	gorobiru	3	L	ru	1	D	D
Medium	keleku	kelekuru	kerekuru	3	L	ru	2	D	D
Medium	korogo	kologoli	korogoli	3	R	li	1	D	H
Medium	nirobu	nilobuli	nirobuli	3	R	li	2	D	H
Medium	nereme	nelemeru	neremeru	3	R	ru	1	H	D
Medium	kuroni	kuloniru	kuroniru	3	R	ru	2	H	D
Medium	tuluge	tulugeli	turugeli	4	L	li	1	H	H
Medium	noledu	noleduli	noreduli	4	L	li	2	H	H
Medium	deludo	deludoru	derudoru	4	L	ru	1	D	D
Medium	bilegi	bilegiru	biregiru	4	L	ru	2	D	D
Medium	tiriti	tilitili	tiritili	4	R	li	1	D	H
Medium	gerude	geludeli	gerudeli	4	R	li	2	D	H
Medium	duribo	duliboru	duriboru	4	R	ru	1	H	D
Medium	borugu	boluguru	boruguru	4	R	ru	2	H	D
Long	letubi	letubili	retubili	1	L	li	1	H	H
Long	linode	linodeli	rinodeli	1	L	li	2	H	H
Long	lunedo	lunedoru	runedoru	1	L	ru	1	D	D
Long	lotiku	lotikuru	rotikuru	1	L	ru	2	D	D
Long	rupimu	lupimuli	rupimuli	1	R	li	1	D	H
Long	ronupe	lonupeli	ronupeli	1	R	li	2	D	H
Long	regedi	legediru	regediru	1	R	ru	1	H	D
Long	ritoko	litokoru	ritokoru	1	R	ru	2	H	D
Long	limegu	limeguli	rimeguli	2	L	li	1	H	H
Long	lugupi	lugupili	rugupili	2	L	li	2	H	H
Long	lokite	lokiteru	rokiteru	2	L	ru	1	D	D
Long	lemogo	lemogoru	remogoru	2	L	ru	2	D	D
Long	romuge	lomugeli	romugeli	2	R	li	1	D	H
Long	rebeti	lebetili	rebetili	2	R	li	2	D	H
Long	ribopo	liboporu	riboporu	2	R	ru	1	H	D
Long	rumibu	lumiburu	rumiburu	2	R	ru	2	H	D
Long	ledimo	ledimoli	redimoli	3	L	li	1	H	H
Long	lokenu	lokenuli	rokenuli	3	L	li	2	H	H
Long	lugoni	lugoniru	rugoniru	3	L	ru	1	D	D

Continued on next page

Table A4 – Testing stimuli (*Continued from previous page*)

Distance	Stem	Option 1	Option 2	Speaker	Liquid	Suffix	Block	Faith	First
Long	lipube	lipuberu	ripuberu	3	L	ru	2	D	D
Long	ruteki	lutekili	rutekili	3	R	li	1	D	H
Long	rikono	likonoli	rikonoli	3	R	li	2	D	H
Long	renitu	lenituru	renituru	3	R	ru	1	H	D
Long	rodume	lodumeru	rodumeru	3	R	ru	2	H	D
Long	lipoke	lipokeli	ripokeli	4	L	li	1	H	H
Long	lebito	lebitoli	rebitoli	4	L	li	2	H	H
Long	lobupu	lobupuru	robupuru	4	L	ru	1	D	D
Long	ludemi	ludemiru	rudemiru	4	L	ru	2	D	D
Long	rodobo	lodoboli	rodoboli	4	R	li	1	D	H
Long	rugidu	lugiduli	rugiduli	4	R	li	2	D	H
Long	rekune	lekuneru	rekuneru	4	R	ru	1	H	D
Long	ripegi	lipegiru	ripegiru	4	R	ru	2	H	D

A3. Statistical models

Here we present full details on the mixed-effects logistic regression models fitted to the responses from each experiment, as reported in the main text. Note that for each of Experiments 1 and 2, two different sets of data were analyzed: the responses from the first 16 participants within each condition, and the responses from the first 12 'successful learners' in each of the experimental groups (compared against the first 12 participants in the Control group). See the main text for details on the criteria used for successful learning.

To minimize issues of collinearity, all predictors were converted into numerical variables, centered around their grand mean, before any model was fitted. For the three-level Distance factor, this entailed creating three such centered predictors, for Short-, Medium-, and Long-range, respectively. In the fitted models reported below, only the latter two were included as predictors, in effect defining Short-range as the baseline (reference) level of the Distance factor. As detailed in Section 3.3.1, each of these models was subsequently re-fitted with Short-range substituted for Medium- and Long-range, respectively, so as to obtain the desired odds ratios (and significance values) for the situation where Medium-range or Long-range served as the baseline reference.

For each data set analyzed, the appropriate structure of the random-effects component of the logit mixed model was determined by the procedure recommended by Bates, Kliegl, Vasishth, and Baayen (2015). The goal of this procedure was to adopt a model that has as rich a random-effects structure as allowed by the data, without resulting in an overparametrised model. Tools from the *RePsychLing* package accompanying Bates et al. (2015) were employed to assess overidentification in candidate models (by Principal Components Analysis of the distribution of variance components in the model) to test whether the model matrix was degenerate. The general process was as follows. First a model with maximal random-effects structure was fitted to the data; 'maximal' here entails random intercepts, slopes, and interaction terms for both subjects and items, as well as random correlation parameters between all pairs of by-subject terms and similarly for the by-item terms. Inevitably, such a 'full' model was always massively overidentified (with no fewer than 82 random-effects parameters on top of the 13 fixed-effect terms). However, the full model provided a baseline for goodness-of-fit comparison for all subsequent (reduced) models; for each alternative model considered, it was rejected if a likelihood-ratio test (with the `anova` function in *R*) showed it to be a significantly worse fit than the original full model. The next step was to fit a corresponding model but with no random correlation parameters included. (This occasionally resulted in a significant loss in fit relative to the full model, indicating that some correlation parameters would likely need to be re-introduced later in the process, as described below.) In all cases, this no-correlations counterpart of the full model was also overidentified.

Individual variance components were then pruned from the model one by one, starting with those with the lowest variance estimates (though respecting the general principle of always removing a higher-order interaction component before any lower-order term nested under it). For each dropped variance component, the goodness-of-fit of the resulting model was compared against both its counterpart with the component included and the original full model. This simplification process was halted once a non-degenerate (fully identified) model was reached, from which no additional random-effect terms could be dropped with-

out significant loss of fit. At this point, correlation parameters for all the random-effect terms remaining in the model (intercepts, slopes, interactions) were reintroduced. In all cases this resulted in a significantly better fitting, but degenerate (non-identified) model. Correlation parameters were then dropped one by one, targeting ones with relatively low estimate (absolute) values, until the model was no longer degenerate; this usually left only a very small number of correlation terms (three or fewer) in the model. As a final check, the goodness of fit of the resulting model was compared against the corresponding model with no correlation parameters as well as the original full model.

For all four experiments, the procedure just described was followed in the analysis of the first 16 participants per condition. Recall that for Experiments 1 and 2, follow-up analyses were carried out on the first 12 participants in each experimental group who successfully learned the pattern they had been trained on (with comparison of these against the first 12 participants recruited in the control group). For comparison purposes, the random-effects structure from the analysis of the corresponding first-16-participants data was carried over to this follow-up analysis of the first-12-learners data.

Details for the (final) logit mixed models for all four experiments are provided in Tables A5 to A10. Note that for Experiments 1 and 3, where the experimental groups were exposed to a pattern of liquid harmony (Tables A5, A6, and A9), the dependent variable was selection of the affixed form with *identical* (agreeing, harmonic) liquids in stem and suffix on a given trial, e.g., choosing [dilopo-li] over [diropo-li], or [diropo-ru] over [dilopo-ru], for an unaffixed stem [dilopo]. For Experiments 2 and 4, which involved exposure to a pattern of liquid dissimilation (Tables A7, A8, and A10), the dependent variable was the exact inverse of this: selection of the affixed form with *non-identical* (disagreeing, disharmonic) liquids in stem and suffix.

Table A5: Full summary of the mixed-effects logistic regression model for Experiment 1 (first 16 participants per condition).

Harmony ∼ HarmonySecond + HarmonyFaithful + MediumRange + LongRange + S-Harm + M-Harm + MediumRange:S-Harm + MediumRange:M-Harm + LongRange:S-Harm + LongRange:M-Harm + (1 | Item) + (1 + LongRange | Subject) + (0 + HarmonySecond + LongRange:HarmonySecond | Subject) + (0 + HarmonyFaithful | Subject) + (0 + LongRange:HarmonyFaithful | Subject)			
$N_{observations}$	4518	Log-likelihood	−2184.2
N_{items}	96	AIC	4408.4
$N_{subjects}$	48	BIC	4536.8

Random Effects

Effect | Grouping Variable	SD (σ)
Intercept | Item	0.3148
Intercept | Subject	0.7483
Harmony Second | Subject	0.7206
Harmony Faithful | Subject	1.9585
Long-range | Subject	0.6884
Long-range × Harmony Second | Subject	0.7381
Long-range × Harmony Faithful | Subject	1.5185
Correlation Parameter	r
(Intercept | Subj. ∼ (Long-range | Subj.)	−0.90
(Harmony Second | Subj.) ∼ (Long-range × Harmony Second | Subj.)	0.70

Fixed Effects

Coefficient	Estimate	SE	$Pr(>\|z\|)$
Intercept	0.43987	0.12132	0.0003
Harmony Second	−0.51084	0.14367	0.0004
Harmony Faithful	2.61950	0.30531	< 0.0001
Medium-range	0.08113	0.18412	0.6595
Long-range	−0.16051	0.25510	0.5292
S-Harm	1.30893	0.38182	0.0006
M-Harm	1.25831	0.38503	0.0011
Medium-range × S-Harm	−1.00769	0.23162	< 0.0001
Long-range × S-Harm	−0.95505	0.34226	0.0053
Medium-range × M-Harm	0.28217	0.23734	0.2345
Long-range × M-Harm	−0.38150	0.34577	0.2699

Table A6: Full summary of the mixed-effects logistic regression model for Experiment 1 (first 12 successful learners per experimental condition + first 12 Control subjects).

Harmony ~ HarmonySecond + HarmonyFaithful + MediumRange + LongRange + S-Harm + M-Harm + MediumRange:S-Harm + MediumRange:M-Harm + LongRange:S-Harm + LongRange:M-Harm + (1 | Item) + (1 + LongRange | Subject) + (0 + HarmonySecond + LongRange:HarmonySecond | Subject) + (0 + HarmonyFaithful | Subject) + (0 + LongRange:HarmonyFaithful | Subject)

$N_{observations}$	3400	Log-likelihood	−1477.2
N_{items}	96	AIC	2994.3
$N_{subjects}$	36	BIC	3117.0

Random Effects

Effect | Grouping Variable	SD (σ)
Intercept | Item	0.2920
Intercept | Subject	0.3881
Harmony Second | Subject	0.4058
Harmony Faithful | Subject	1.8905
Long-range | Subject	0.7074
Long-range × Harmony Second | Subject	0.8100
Long-range × Harmony Faithful | Subject	1.2650
Correlation Parameter	r
(Intercept | Subj.) ~ (Long-range | Subj.)	−0.44
(Harmony Second | Subj.) ~ (Long-range × Harmony Second | Subj.)	0.31

Fixed Effects

Coefficient	Estimate	SE	$Pr(>\|z\|)$
Intercept	1.00670	0.09180	< 0.0001
Harmony Second	−0.37010	0.13414	0.0058
Harmony Faithful	2.55813	0.34168	< 0.0001
Medium-range	0.04296	0.20076	0.8305
Long-range	−0.19051	0.29228	0.5145
S-Harm	2.60083	0.32002	< 0.0001
M-Harm	2.55636	0.32189	< 0.0001
Medium-range × S-Harm	−2.06848	0.28971	< 0.0001
Long-range × S-Harm	−2.19245	0.42101	< 0.0001
Medium-range × M-Harm	0.12619	0.30437	0.6784
Long-range × M-Harm	−1.08123	0.42259	0.0105

Table A7: Full summary of the mixed-effects logistic regression model for Experiment 2 (first 16 participants per condition).

Dissimilation ~ DissimilationSecond + DissimilationFaithful + MediumRange + LongRange + S-Diss + M-Diss + MediumRange:S-Diss + MediumRange:M-Diss + LongRange:S-Diss + LongRange:M-Diss + (1 | Item) + (1 | Subject) + (0 + DissimilationSecond | Subject) + (0 + DissimilationFaithful | Subject) + (0 + MediumRange | Subject) + (0 + LongRange | Subject) + (0 + MediumRange:DissimilationFaithful + LongRange:DissimilationFaithful | Subject)

$N_{observations}$	4534	Log-likelihood	−2102.2
N_{items}	96	AIC	4244.3
$N_{subjects}$	48	BIC	4372.7

Random Effects

Effect | Grouping Variable	*SD* (σ)
Intercept | Item	0.3274
Intercept | Subject	0.7053
Dissimilation Second | Subject	0.8077
Dissimilation Faithful | Subject	1.9942
Medium-range | Subject	0.3221
Long-range | Subject	0.3313
Medium-range × Dissimilation Faithful | Subject	1.5013
Long-range × Dissimilation Faithful | Subject	1.7347
Correlation Parameter	*r*
(Medium-range × Diss. Faithf. | Subj.) ~ (Long-range × Diss. Faithf. | Subj.)	0.67

Fixed Effects

Coefficient	Estimate	*SE*	$Pr(>\|z\|)$
Intercept	0.84548	0.11733	< 0.0001
Dissimilation Second	−0.76281	0.15987	< 0.0001
Dissimilation Faithful	2.43353	0.31083	< 0.0001
Medium-range	−0.11150	0.20814	0.5922
Long-range	0.21218	0.20703	0.3054
S-Diss	2.31047	0.32805	< 0.0001
M-Diss	1.43201	0.31669	< 0.0001
Medium-range × S-Diss	−1.85678	0.28549	< 0.0001
Long-range × S-Diss	−2.56493	0.29045	< 0.0001
Medium-range × M-Diss	−0.02677	0.28491	0.9251
Long-range × M-Diss	−1.04038	0.27559	0.0002

Table A8: Full summary of the mixed-effects logistic regression model for Experiment 2 (first 12 successful learners per experimental condition + first 12 Control subjects).

Dissimilation ~ DissimilationSecond + DissimilationFaithful + MediumRange + LongRange + S-Diss + M-Diss + MediumRange:S-Diss + MediumRange:M-Diss + LongRange:S-Diss + LongRange:M-Diss + (1 | Item) + (1 | Subject) + (0 + DissimilationSecond | Subject) + (0 + DissimilationFaithful | Subject) + (0 + MediumRange | Subject) + (0 + LongRange | Subject) + (0 + MediumRange:DissimilationFaithful + LongRange:DissimilationFaithful | Subject)

$N_{observations}$	3403	Log-likelihood	−1550.8
N_{items}	96	AIC	3141.6
$N_{subjects}$	36	BIC	3264.3

Random Effects

Effect | Grouping Variable	*SD* (σ)
Intercept | Item	0.2795
Intercept | Subject	0.4829
Dissimilation Second | Subject	0.4923
Dissimilation Faithful | Subject	1.9479
Medium-range | Subject	< 0.0001
Long-range | Subject	0.1994
Medium-range × Dissimilation Faithful | Subject	1.5099
Long-range × Dissimilation Faithful | Subject	1.7106
Correlation Parameter	*r*
(Medium-range × Diss. Faithf. | Subj.) ~ (Long-range × Diss. Faithf. | Subj.)	0.67

Fixed Effects

Coefficient	Estimate	*SE*	$Pr(>\|z\|)$
Intercept	1.01268	0.10235	< 0.0001
Dissimilation Second	−0.57817	0.13990	< 0.0001
Dissimilation Faithful	2.15613	0.34737	< 0.0001
Medium-range	−0.03413	0.20099	0.8652
Long-range	0.19720	0.20875	0.3448
S-Diss	2.01229	0.29585	< 0.0001
M-Diss	2.25555	0.30683	< 0.0001
Medium-range × S-Diss	−1.70275	0.28233	< 0.0001
Long-range × S-Diss	−2.21943	0.29731	< 0.0001
Medium-range × M-Diss	0.34803	0.32061	0.2778
Long-range × M-Diss	−1.53086	0.30706	< 0.0001

Table A9: Full summary of the mixed-effects logistic regression model for Experiment 3 (first 16 participants per condition).

Harmony $\sim$ HarmonySecond + HarmonyFaithful + MediumRange + LongRange + S-Harm-M-Faith + M-Harm-S-Faith + MediumRange:S-Harm-M-Faith + MediumRange:M-Harm-S-Faith + LongRange:S-Harm-M-Faith + LongRange:M-Harm-S-Faith + (1 | Item) + (1 + LongRange | Subject) + (0 + HarmonySecond + LongRange:HarmonySecond | Subject) + (0 + HarmonyFaithful | Subject) + (0 + LongRange:HarmonyFaithful | Subject)			
$N_{observations}$	4490	Log-likelihood	-2101.2
N_{items}	96	AIC	4242.3
$N_{subjects}$	48	BIC	4370.5

Random Effects

Effect | Grouping Variable	SD (σ)
Intercept | Item	0.3245
Intercept | Subject	0.5633
Harmony Second | Subject	0.6900
Harmony Faithful | Subject	2.1714
Long-range | Subject	0.05488
Long-range $\times$ Harmony Faithful | Subject	1.59367
Correlation Parameter	r
(Intercept | Subj. $\sim$ (Harmony Second | Subj.)	0.62
(Intercept | Subj. $\sim$ (Harmony Faithful | Subj.)	-0.55
(Harmony Second | Subj.) $\sim$ (Harmony Faithful | Subj.)	-0.19

Fixed Effects

Coefficient	Estimate	SE	$Pr(>\|z\|)$
Intercept	0.37350	0.09907	0.0002
Harmony Second	-0.45671	0.14851	0.0021
Harmony Faithful	2.78418	0.33468	< 0.0001
Medium-range	0.09294	0.18578	0.6169
Long-range	-0.19822	0.18699	0.2891
S-Harm-M-Faith	1.96544	0.24315	< 0.0001
M-Harm-S-Faith	0.78951	0.22659	0.0005
Medium-range $\times$ S-Harm-M-Faith	-1.53211	0.24457	< 0.0001
Long-range $\times$ S-Harm-M-Faith	-1.53562	0.25257	< 0.0001
Medium-range $\times$ M-Harm-S-Faith	0.11878	0.23573	0.6144
Long-range $\times$ M-Harm-S-Faith	-0.16419	0.23987	0.4937

Table A10: Full summary of the mixed-effects logistic regression model for Experiment 4 (first 16 participants per condition).

Dissimilation ~ DissimilationSecond + DissimilationFaithful + MediumRange + LongRange + S-Diss-M-Faith + M-Diss-S-Faith + MediumRange:S-Diss-M-Faith + MediumRange:M-Diss-S-Faith + LongRange:S-Diss-M-Faith + LongRange:M-Diss-S-Faith + (1 | Item) + (1 | Subject) + (0 + DissimilationSecond | Subject) + (0 + DissimilationFaithful | Subject) + (0 + MediumRange | Subject) + (0 + LongRange | Subject) + (0 + MediumRange:DissimilationFaithful + LongRange:DissimilationFaithful | Subject)

$N_{observations}$	4518	Log-likelihood	−2221.8
N_{items}	96	AIC	4487.6
$N_{subjects}$	48	BIC	4628.8

Random Effects

Effect | Grouping Variable	SD (σ)
Intercept | Item	0.2652
Intercept | Subject	0.3767
Dissimilation Second | Subject	0.6548
Dissimilation Faithful | Subject	1.9440
Medium-range | Subject	0.6700
Long-range | Subject	0.5657
Medium-range × Dissimilation Faithful | Subject	1.1381
Long-range × Dissimilation Faithful | Subject	1.0173
Correlation Parameter	r
(Intercept | Subj.) ~ (Dissimilation Faithful | Subj.)	0.60
(Medium-range | Subj.) ~ (Long-range | Subj.)	0.89
(Medium-range × Diss. Faithf. | Subj.) ~ (Long-range × Diss. Faithf. | Subj.)	−0.90

Fixed Effects

Coefficient	Estimate	SE	$Pr(>\|z\|)$
Intercept	0.47786	0.07546	< 0.0001
Dissimilation Second	−0.46044	0.13429	0.0006
Dissimilation Faithful	2.71046	0.30085	< 0.0001
Medium-range	−0.15428	0.25325	0.5424
Long-range	0.15468	0.23290	0.5066
S-Diss-M-Faith	2.03793	0.26494	< 0.0001
M-Diss-S-Faith	0.04013	0.24261	0.8686
Medium-range × S-Diss-M-Faith	−1.75389	0.35343	< 0.0001
Long-range × S-Diss-M-Faith	−2.14860	0.33003	< 0.0001
Medium-range × M-Diss-S-Faith	0.55088	0.33554	0.1006
Long-range × M-Diss-S-Faith	−0.37554	0.31020	0.2260

A4. Experiments 3 and 4: Individual results

Figures A1 and A2 show the results for individual participants in the two training conditions that conformed to a *strictly beyond-transvocalic* locality pattern: the M-Harm-S-Faith condition of Experiment 3 and the M-Diss-S-Faith condition of Experiment 4, respectively. For the latter, these represent the original pool of 16 participants; the additional 24 recruits are not shown (but see Figure 5 in the main paper, which includes those five participants in that additional group who did display a preference for dissimilation in Medium-range contexts).

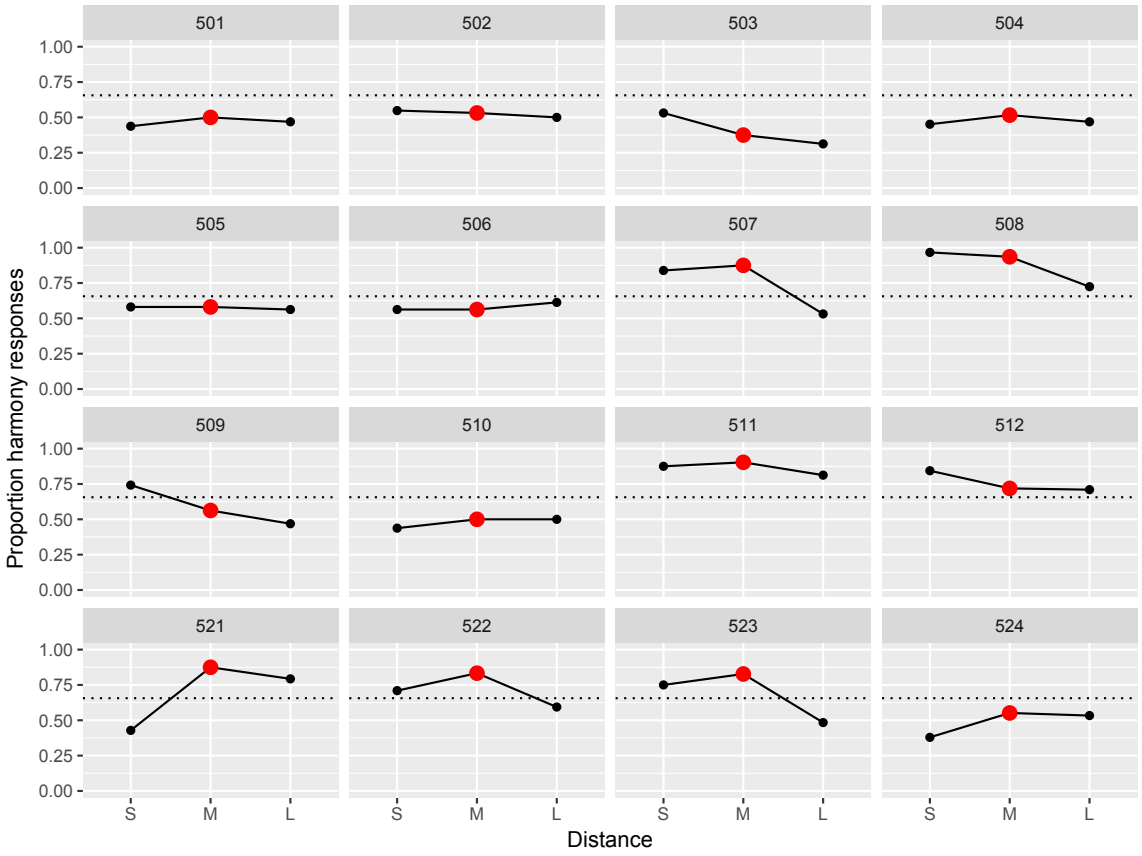

Figure A1: Experiment 3: Individual results for M-Harm-S-Faith group. Performance in the context where harmony was encountered in training is highlighted in red. The dotted horizontal line indicates the threshold criterion for reliable application of harmony (95% confidence limit on one-tailed binomial test).

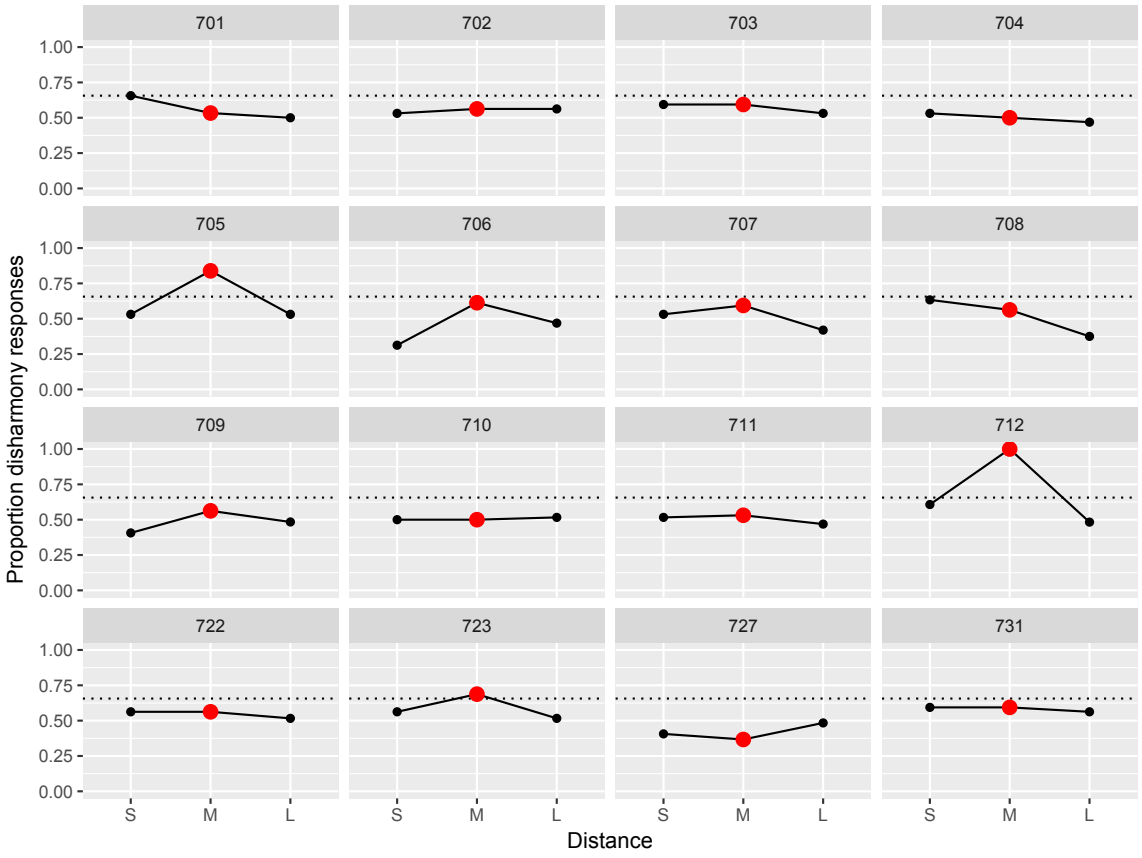

Figure A2: Experiment 4: Individual results for M-Diss-S-Faith group (original pool of 16 participants). Performance in the context where dissimilation was encountered in training is highlighted in red. The dotted horizontal line indicates the threshold criterion for reliable application of dissimilation (95% confidence limit on one-tailed binomial test).

A5. Factorial typology of ABC model

Here we describe in greater detail the factorial typology of locality patterns for liquid-liquid dependencies that is generated by the Agreement by Correspondence model (Bennett, 2015; Hansson, 2010; Rose & Walker, 2004), as described in Sections 1.2.2, 1.3, and 8.2.1.

A5.1. Constraint set

The full set of constraints that form the basis of our typology is laid out in (1)–(3). Two comments on this constraint set are in order. First, in (1) we use an *Input-Output* faithfulness constraint as the force counteracting any restrictions on liquid-liquid co-occurrence. However, since our experimental stimuli all involved morpho-phonological alternations between a bare stem and an affixed form, an *Output-Output* faithfulness constraint (OO-IDENT[lateral]) would have the same effect. In the ranking configurations described in (4) below, any constraint that crucially dominates IO-IDENT[lateral] will also need to dominate OO-IDENT[lateral] in order for the dependency pattern in question to hold in affixed forms like those in our stimuli. When discussing learning biases in the form of initial ranking (or weighting) of constraints in Section 9.2.1 of the article, we focus on Output-Output rather than Input-Output faithfulness, as there are strong reasons to assume that the former, rather than the latter, is what is favoured in the learner's initial state (see note 9 in the main article and references cited therein).

Second, as explained in Section 1.3 of the article, the 'constraint' CC-NOCORR in (3c) is merely a shorthand placeholder for any constraint, or set of constraints, whose *effect* is to prohibit correspondence for all co-occurring C…C segment pairs, regardless of locality, featural content, or the like. For the purpose of our experiments, where the co-occurring liquids under consideration are always heteromorphemic (stem versus suffix), the constraint CC-EDGE(morpheme) (Bennett, 2015), which requires correspondents to be *tautomorphemic* could play this role. That constraint would be violated by any output form of the type $[\ldots\{r,l\}_x\ldots]_{Stem} + [\ldots\{r,l\}_x\ldots]_{Suffix}$, regardless of which liquids are involved, or how far apart they are from each other. As explained in the article (Section 1.3), an equivalent but morphology-blind ban on correspondence could result from the combined and mutually exclusive demands of the constraints CC-EDGE(σ) (correspondents must be tautosyllabic) and CC-SROLE (correspondents must occupy matching syllable constituents).

(1) Faithfulness constraint:

IO-IDENT[lateral]
An output segment must preserve the [±lateral] value of its input correspondent. (No change to underlying [±lateral] specifications.)

(2) CORR-X constraints, imposing correspondence on similar segments:

a. CORR-LIQ
Co-occurring liquids must stand in correspondence. (Correspondence required in {l…r, r…l, l…l, r…r}.)

b. CORR-LIQ$_{CVC}$
Co-occurring liquids, if separated by at most a vowel, must stand in correspondence. (Correspondence required in {lVr, rVl, lVl, rVr}.)

c. CORR-LIQ[αlateral]
Co-occurring identical liquids must stand in correspondence. (Correspondence required in {l…l, r…r}.)

d. CORR-LIQ[αlateral]$_{CVC}$
Co-occurring identical liquids, if separated by at most a vowel, must stand in correspondence. (Correspondence required in {lVl, rVr}.)

(3) 'CC-Limiter' constraints, placing restrictions on correspondents:

a. PROXIMITY (a.k.a. CC-SYLLADJ)
Correspondent segments must be tautosyllabic or in adjacent syllables. (No correspondence across an intervening syllable.)

b. CC-IDENT[lateral]
Correspondent segments must have agreeing specifications for [±lateral]. (No correspondence between disagreeing liquids.)

c. 'CC-NOCORR' (see clarification in text)
Correspondence is not allowed, regardless of segment type, position, or distance. (No correspondence between any co-occurring output segments.)

The constraints in (1), (2a), (2c), and (3a) all figure in the discussion in Section 1.3. The locality-restricted 'CVC' versions of the CORR-X constraints in (2a), (2c), defined in (2b) and (2d), respectively, are essential in order for the generated typology to include *transvocalic dissimilation* (i.e., dissimilation confined to CVC contexts). Such constraints were originally proposed in Hansson (2001), as well as in a 2001 manuscript version of Rose and Walker (2004), to account for *transvocalic harmony*. Rose and Walker (2004) instead propose the PROXIMITY constraint in (3a) for that purpose. Bennett (2015) includes both constraint types (slightly refining the formal definition of the latter, and renaming it CC-SYLLADJ).

The constraint CC-IDENT[lateral] in (3b), which requires agreement in [±lateral] between correspondent liquids, is essential for deriving *liquid harmony* (together with one or both of the CORR-LIQ constraints in (2a)–(2b), which force liquids into correspondence). The shorthand constraint CC-NOCORR in (3c), which is effectively a context-free ban on correspondence altogether (see above and Section 1.3 of the article), is essential for deriving *unbounded dissimilation*.

A5.2. Generating the typology

The full set of eight constraints in (1)–(3) defines 40,320 (8!, eight factorial) possible ranking permutations. As mentioned in Section 8.2.1, we used the *OT-Help 2.0* software package (Staubs et al., 2010) to determine the range of distinct *languages* (patterns) that these represent. Three different source files for the OT-Help calculations are included in the supplementary materials. The file `ABCDissHarm_Full.txt` uses the full eight-constraint set. We also calculated typologies for two reduced versions of the constraint set, excluding either the 'CVC' versions of the two CORR-X constraints, (2b) and (2d), or the

PROXIMITY constraint, (3a); the source files for these are `ABCDissHarm_NoCVC.txt` and `ABCDissHarm_NoProxim.txt`, respectively. In the OT-Help source files, the abbreviated constraint labels should be self-explanatory, except perhaps that `CorrLiq`, `CorrLiqCVC`, `CorrSame`, and `CorrSameCVC` stand for the four CORR-X constraints in (2a)–(2d), in that order.

Each source file contains tableaux for eight different input representations. These all have the structure CVCVCV, four with a liquid pair in C_2–C_3 (transvocalic distance) and four with a liquid pair in C_1–C_3 (beyond-transvocalic distance). At each of these two distances, there is one input (tableau) for each of the four logically possible liquid-liquid combinations: /l…l/, /r…r/, /r…l/, and /l…r/.

For each of these eight inputs (4 liquid combinations $\times$ 2 distances), the relevant tableau provides four candidate output representations, which are intended to mimic the 2AFC response options in the experiments. Two of the candidates preserve unchanged the liquid-liquid sequence of the input; in one candidate, the two liquids stand in a correspondence relation, in the other they do not. In the other two candidates, the first of the two liquids is altered relative to the input; again, the two liquids are either correspondents of each other or not. Thus, for example, the candidate set for both the input /CVlVlV/ and the input /CVrVlV/ was {[CVl$_x$Vl$_y$V], [CVl$_x$Vl$_x$V], [CVr$_x$Vl$_y$V], [CVr$_x$Vl$_x$V]}. In the source files, uppercase is used to indicate correspondence; e.g., [CVl$_x$Vl$_y$V] and [CVl$_x$Vl$_x$V] are rendered as `cvlvlv` and `cvLvLv`, respectively.

Given the sizes of the constraint sets used, the total numbers of possible ranking permutations are $8! = 40,320$ (`ABCDissHarm_Full.txt`), $7! = 5,040$ (`ABCDissHarm_Full.txt`), and $6! = 720$ (`ABCDissHarm_NoCVC.txt`). In terms of the combinations of input-output mappings generated across the eight tableaux provided, the number of distinct 'languages' generated is much smaller: 14, 13, and 12, respectively. However, many of these are non-distinct from each other in terms of the *phonetic output*, in that they differ from each other solely in the covert correspondence relations assigned to particular outputs. To take a simple example, a full half of the 12 different 'languages' generated by the smallest constraint set (`ABCDissHarm_NoCVC.txt`) yield a fully faithful output for all eight inputs; in other words, they all represent the complete absence of any co-occurrence restrictions on liquid…liquid sequences. These six 'languages,' listed in Table A11, are all phonetically indistinguishable; they differ only in whether an abstract structural relation of correspondence is present for all pairs of co-occurring liquids (a), none at all (f), or for some pairs but not others: transvocalic and/or identical pairs (b), identical pairs only (c), transvocalic pairs only (d), or pairs that are simultaneously transvocalic and identical (e). To reduce clutter, we only show subscript indices where correspondence is present, and no indices where the two liquids are not in correspondence.

When such observationally equivalent 'languages' are conflated, we arrive at the much smaller typology of truly distinct liquid-liquid dependency patterns (or lack thereof) reported below and referenced in Sections 1.3 and 9.2.1 of the main article.

Table A11: Sets of input-output mappings in six no-dependency 'languages'

	/CVlVlV/	/lVCVlV/	/CVrVrV/	/rVCVrV/	/CVrVlV/	/rVCVlV/	/CVlVrV/	/lVCVrV/
a.	CVl_xVl_xV	l_xVCVl_xV	CVr_xVr_xV	r_xVCVr_xV	CVr_xVl_xV	r_xVCVl_xV	CVl_xVr_xV	l_xVCVr_xV
b.	CVl_xVl_xV	l_xVCVl_xV	CVr_xVr_xV	r_xVCVr_xV	CVr_xVl_xV	rVCVlV	CVl_xVr_xV	lVCVrV
c.	CVl_xVl_xV	l_xVCVl_xV	CVr_xVr_xV	r_xVCVr_xV	CVrVlV	rVCVlV	CVlVrV	lVCVrV
d.	CVl_xVl_xV	lVCVlV	CVr_xVr_xV	rVCVrV	CVr_xVl_xV	rVCVlV	CVl_xVr_xV	lVCVrV
e.	CVl_xVl_xV	lVCVlV	CVr_xVr_xV	rVCVrV	CVrVlV	rVCVlV	CVlVrV	lVCVrV
f.	CVlVlV	lVCVlV	CVrVrV	rVCVrV	CVrVlV	rVCVlV	CVlVrV	lVCVrV

A5.3. Resulting typology

The typology generated by the full constraint set is shown in (4); this is the basis for the list of dependency types given in Section 9.2.1 of the article.[1] Included under each type in (4) is a description of any sets of constraint-ranking relations that are crucial for generating the pattern in question. In some cases, multiple distinct (i.e., potentially incompatible) sets of ranking configurations can result in the same pattern; this is indicated with an 'either–or' disjunction.

(4) Factorial typology generated by the ABC constraint set

a. Unbounded harmony

i. CORR-LIQ $\gg$ { CC-EDGE(morph), PROXIMITY, IO-IDENT[lat] }

ii. CC-IDENT[lat] $\gg$ IO-IDENT[lat]

b. Unbounded dissimilation

i. CC-EDGE(morph) $\gg$ { CORR-LIQ, $\text{CORR-LIQ}_{\text{CVC}}$, IO-IDENT[lat] }

ii. CORR-LIQ[αlat] $\gg$ IO-IDENT[lat]

c. Strictly transvocalic harmony

either

i. $\text{CORR-LIQ}_{\text{CVC}}$ $\gg$ CC-EDGE(morph) $\gg$ { CORR-LIQ, CORR-LIQ[αlat] }

ii. { $\text{CORR-LIQ}_{\text{CVC}}$, CC-IDENT[lat] } $\gg$ IO-IDENT[lat] $\gg$ CORR-LIQ[αlat]

or

i. PROXIMITY $\gg$ CORR-LIQ $\gg$ { CC-EDGE(morph), IO-IDENT[lat] }

ii. CC-IDENT[lat] $\gg$ IO-IDENT[lat] $\gg$ CORR-LIQ[αlat]

d. Strictly transvocalic dissimilation

i. CC-EDGE(morph) $\gg$ { CORR-LIQ, $\text{CORR-LIQ}_{\text{CVC}}$, IO-IDENT[lat] }

ii. $\text{CORR-LIQ}[\alpha\text{lat}]_{\text{CVC}}$ $\gg$ IO-IDENT[lat] $\gg$ CORR-LIQ[αlat]

[1] *OT-Help* calculates typologies not only based on constraint ranking (strict domination), as in Optimality Theory, but also with weighted constraints, as in Harmonic Grammar (Pater, 2009; Potts, Pater, Jesney, Bhatt, & Becker, 2010). For the constraint set in (1)–(3), both options yield an identical range of distinct dependency patterns; the same remains true under the reduced constraint sets without either PROXIMITY or the two 'CVC' variants of CORR-X constraints. The points made in Section 9.2.1 regarding learning biases (in the learner's initial state) can thus be framed interchangeably in terms of ranking or weighting of constraints.

e. Strictly beyond-transvocalic dissimilation
either
i. CORR-LIQ$_{CVC}$ $\gg$ CC-EDGE(morph) $\gg$ { CORR-LIQ, IO-IDENT[lat] }
ii. CORR-LIQ[αlat] $\gg$ IO-IDENT[lat] $\gg$ CC-IDENT[lat]
or
i. PROXIMITY $\gg$ CORR-LIQ $\gg$ { CC-EDGE(morph), CC-IDENT[lat] }
ii. CORR-LIQ[αlat] $\gg$ IO-IDENT[lat] $\gg$ CC-IDENT[lat]

f. Transvocalic harmony with beyond-transvocalic dissimilation
either
i. CORR-LIQ$_{CVC}$ $\gg$ CC-EDGE(morph) $\gg$ { CORR-LIQ, IO-IDENT[lat] }
ii. { CORR-LIQ[αlat], CC-IDENT[lat] } $\gg$ IO-IDENT[lat]
or
i. PROXIMITY $\gg$ CORR-LIQ $\gg$ { CC-EDGE(morph), IO-IDENT[lat] }
ii. { CORR-LIQ[αlat], CC-IDENT[lat] } $\gg$ IO-IDENT[lat]

g. No dependency
(all ranking permutations not falling under any of the above)

When PROXIMITY is left out of the constraint set, the full range of patterns remains the same; the sole difference is that only one of the alternative ranking configurations remains available for each of (4c), (4e), and (4f). On the other hand, when the 'CVC' versions of the two CORR-X constraints are eliminated, the strictly transvocalic dissimilation option in (4d) disappears from the typology; this is because that pattern crucially relies on CORR-LIQ$_{CVC}$ being ranked relatively high (crucially dominating input-output faithfulness).

References

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. *ArXiv e-prints*. Retrieved from `https://arxiv.org/abs/1506.04967`

Bennett, W. G. (2015). *The phonology of consonants: Harmony, dissimilation and correspondence*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139683586

Hansson, G. Ó. (2001). *Theoretical and typological issues in consonant harmony* (Doctoral dissertation). University of California, Berkeley.

Hansson, G. Ó. (2010). *Consonant harmony: Long-distance interaction in phonology*. Berkeley, CA: University of California Press. Retrieved from `https://escholarship.org/uc/item/2qs7r1mw`

Pater, J. (2009). Weighted constraints in generative linguistics. *Cognitive Science*, *33*(6), 999–1035. doi: 10.1111/j.1551-6709.2009.01047.x

Potts, C., Pater, J., Jesney, K., Bhatt, R., & Becker, M. (2010). Harmonic grammar with linear programming: From linear systems to linguistic typology. *Phonology*, *27*(1), 77-117.

Rose, S., & Walker, R. (2004). A typology of consonant agreement as correspondence. *Language*, *80*, 475–531. doi: 10.1353/lan.2004.0144

Staubs, R., Becker, M., Potts, C., Pratt, P., McCarthy, J. J., & Pater, J. (2010). *OT-Help 2.0. Software package.* Amherst, MA: University of Massachusetts Amherst. Retrieved from `http://people.umass.edu/othelp/`