

Appendix A

Word ₁	Word ₂	Hybrid Model Prediction	Heuristic Model Prediction	Ground Truth
documentary	series	docuseries, documseries	documeries, documces	docuseries
sea	hijack	seajack	seack	seajack
mom	entrepreneur	momprenneur, momtreprenneur	moreprenneur, momtreprenneur	momtreprenneur
catalog	magazine	catalozine, catalzine	catazine, catalogazine	catazine
japan	english	japelish, japlish	japanglish	janglish

Table 1: Comparison of the heuristic and hybrid models on additional sample inputs

Word ₁	Word ₂	Prediction	Ground Truth
smelly	elephant	smellyphant, smelephant	smellephant
baptist	catholic	baptholic, bapolic	baptilic
meat	potatoes	mettatoes, meatatoes	meatoes
anachronism	acronym	anachronym, anachacronym	anacronym
static	satellite	statellite	statite
boat	hotel	boatel, boattel	botel
float	hotel	flotel	floatel
thumb	type	thypo, thtypo	smellephant
chocolate	milk	chocomilk, chocoilk	chmilk
video	podcast	vodcast, vcat	videocast

Table 2: Sample errors - plausible predictions by the hybrid model. The ground-truth blends botel and floatel show conflict in the choice of structure. Factoring in dominance can resolve the conflict.

Word ₁	Word ₂	Prediction	Ground Truth
crazy	drunk	crzunk, crrunk	crunk
humid	index	humiindex, humindex	humidex
invite	intro	invtro, invintro	invitro
update	insert	updsert, updasert	upsert
magnet	pulsar	maglsar, maulsar	magnetar
bully	suicide	bullyicide, bullice	bullycide
geek	hipster	geekpster, gpster	geekster
prediction	dictionary	predary, preddictionary	predictionary
awkward	fest	awkwfest, awkwest	awkfest
carbon	oxygen	carbxygen, carboxygen	carbogen

Table 3: Word formation errors. These samples violate rules of word formation, hence several blends are not pronounceable. To enforce such rules on the resulting word some form of reinforcement learning has to be applied since gradients cannot be passed back.

Word ₁	Word ₂	Prediction	Ground Truth
spam	blog	spog, spamog	splog
compliment	insult	compsult, complsult	complimentsult
misinformation	entertainment	misintainment, misitianment	misinfotainment
banquet	dance	banance, bannce	banquance
pretend	remember	pretmember, preember	pretember
iodine	ozone	iozone, iodzone	iodozone
market	architecture	marhchitecture, marchitecture	marketecture
fence	hedge	fenedge, feedge	fedge

Table 4: Recognizability errors. In these type of errors the blends 'feel' unnatural and the source words can be difficult to infer

Word ₁	Word ₂	Prediction	Ground Truth
front	trunk	frontrunk, fronrunk	frunk
information	epidemic	infordemic, inforidemic	infodemic
pyke	concrete	pycrete	pycrete
sheldon	amy	sheldamy, shelamy	shemy
arabic	english	arabglish, arablish	arabish

Table 5: Overrepresentation errors. Here the blends tend to keep more source word characters than necessary

Blend Type	Example
Central Placement	<i>advertisement + tease = advertteasement</i>
Prefix Prefix Combination	<i>cybernetic + organism = cyborg</i>
More than 2 sources	<i>thanksgiving + halloween + christmas = thankshallowistmas</i>
Graphemic Changes	<i>fur + explode = fursplode</i>

Table 6: Blends not handled by our system. These blend types had too few examples to sufficiently train out system