Phonology **34** (2017). Supplementary materials

Defying the stimulus: acquisition of complex onsets in Polish

Gaja Jarosz
University of Massachusetts Amherst

Supplementary materials

1 Alternative frequency measures and scales

This section of the supplementary materials considers two alternative ways of analysing the input and the SSP scale respectively.

1.1 Token frequency

For completeness, this section demonstrates that relying on token frequency instead of type frequency does not provide a way out for the lexicalist hypothesis. Figure 7 shows the association between accuracy and token-frequency measures, analogous to the figures above for the type-frequency measures. The results are similar, albeit less promising: token segmental bigram frequency is not positively associated with accuracy.

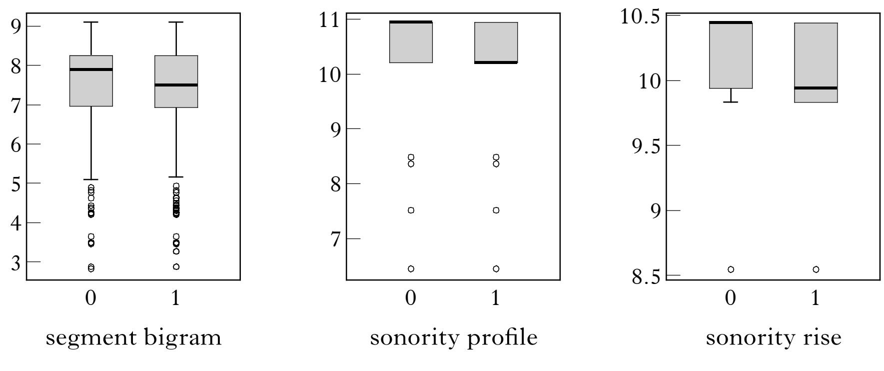

Figure 7
Association between accuracy and token frequency.

A summary of model comparisons analogous to those in Table IV above is shown for token frequencies in Table VI. The pattern of results is similar, except that token segmental bigram frequency is not predictive of production accuracy. Just like the corresponding type frequencies, sonority-profile ($\beta = -0.328$, $z = -4.54$, $p < 0.0001$) and sonority-rise ($\beta = -0.569$, $z = -5.07$, $p < 0.0001$) token frequencies are predictive of production accuracy, but in the wrong direction. The superset models with SSP are superior in all cases, and SSP is retained as a predictor in all but one of 200 backwards elimination bootstrap samples. Thus token frequencies do not provide a way to capture the developmental SSP effect.

		segmental bigram		sonority profile		sonority rise	
	base	+freq	+freq +SSP	+freq	+freq +SSP	+freq	+freq +SSP
D_{xy}	0.425	0.425	0.458	0.445	0.459	0.452	0.458
LR	357.4	357.7	410.2	380.1	412.5	385.6	410.0
$\chi^2(1)$		0.366	52.5	22.7	32.4	28.2	24.4
p		0.545	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Table VI
Summary of statistical tests for type frequencies.

1.2 Finer-grained SSP

The analyses above follow prior modelling studies in assuming the coarse-grained sonority scale that worked well for generating sonority projection in other languages. The granularity of the sonority scale is often debated, however. Could it be that the coarse sonority scale is working against the lexicalist hypothesis by lumping all the obstruents together? This section first demonstrates that the same conclusions about children's sensitivity to the SSP are reached when a finer-grained sonority scale that separates plosives (including affricates) from fricatives is used (Selkirk 1984). Interestingly, the finer-grained sonority scale turns out to be a better predictor of children's production accuracy than the coarse scale. This is unexpected, given that formal analyses of Polish phonology explicitly argue for the coarse-grained scale. It is consistent, however, with recent findings suggesting that sonority-projection effects differentiate fricatives and stops (Tamási & Berent 2014, Lennertz & Berent 2015). The section then shows that the frequency predictions calculated on a finer-grained scale yield qualitatively similar patterns of results to the coarse-grained scale examined above.

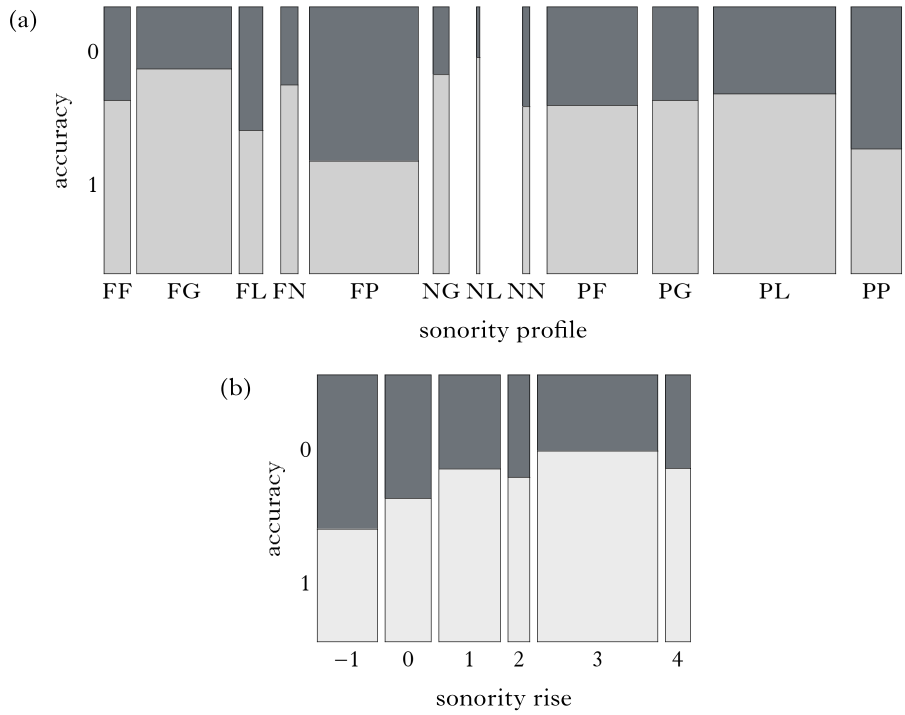

Figure 8

Accuracy by (a) finer-grained sonority profile; (b) finer-grained sonority rise.

Figures 8a and b show the relationship between accuracy and finer-grained sonority profiles and rises respectively (F = fricative, P = plosive). The main differences from before are that finer-grained sonority treats FP as a mild sonority fall, PF as a mild rise like FN, PG as a larger rise than FG or PL, and PL as a larger rise than FL. Inspection of Fig. 8 reveals that most of these difference align well with accuracy: FP is the least accurate cluster type, PF is close in accuracy to FN and children are less accurate on FL than PL. PG does not appear to be favoured by children relative to FG, however. Nonetheless, a nested model comparison shows that a predictor based on the finer-grained SSP (fSSP) is highly significantly predictive of production accuracy after controlling for the various potential confounding variables discussed earlier ($\chi^2(1) = 64.7$; $p < 0.0001$). As expected, the association is positive: higher fSSP is associated with higher accuracy ($\beta = 0.24$, $z = 7.95$, $p < 0.0001$). Out of 200 bootstrap validation samples with backward elimination, fSSP is retained in the model 200 times. Overall, the model shows a small amount of shrinkage: the original D_{xy} is 46.7, the optimism is 0.014 and the corrected D_{xy} is 45.3.

As mentioned earlier, an unexpected finding is that fSSP is a better predictor of children's accuracy than SSP. This is verified by evaluating a superset model that adds SSP. The superset model is not superior to the model with just fSSP ($\chi^2(1) = 0.059$; $p > 0.8$), indicating that SSP is superfluous once fSSP is in the model. Accordingly, the opposite nested model comparison reveals that fSSP makes a significant contribution to a model that already has SSP ($\chi^2(1) = 12.1$; $p < 0.001$). Note that this is not a matter of degrees of freedom, since both SSP and fSSP are continuous predictors with one degree of freedom: this just means that the additional distinctions made on the finer-grained sonority scale are reflected in the children's production accuracy. As further verification of this result, the backwards elimination validation procedure retains fSSP and not SSP.

To conclude this discussion, Table VII shows that calculating frequency along the finer-grained sonority scale does not rescue the lexicalist approach. Just like the corresponding coarse-grained type frequencies, fine-grained sonority-profile ($\beta = -0.143$, $z = -2.32$, $p < 0.05$) and sonority-rise ($\beta = -0.2725$, $z = 2.93$, $p < 0.01$) type frequencies are predictive of production accuracy. The only difference is that sonority-rise frequency is associated in the right direction with accuracy. Just as with coarse-grained frequency, the superset models with SSP are superior in both cases, and SSP is retained as a predictor in all 200 backwards elimination bootstrap samples. (These are the numbers using the weaker SSP predictor – results are similar when fSSP is used.) Thus coarse-grained frequencies do not provide a way to capture the developmental SSP effect.

		sonority profile		sonority rise	
	base	+freq	+freq +SSP	+freq	+freq +SSP
D_{xy}	0.425	0.428	0.471	0.434	0.469
LR	357.4	362.9	429.4	365.9	422.3
$\chi^2(1)$		5.5	66.6	8.5	56.4
p		0.019	<0.0001	0.003	<0.0001

Table VII

Summary of statistical tests for type frequencies using finer-grained sonority scale.

2 Other supplementary materials

Figure 9 presents children's error rates, broken down by sonority profile (see § 3 of the paper), and Table VIII provides a list of Polish initial clusters.

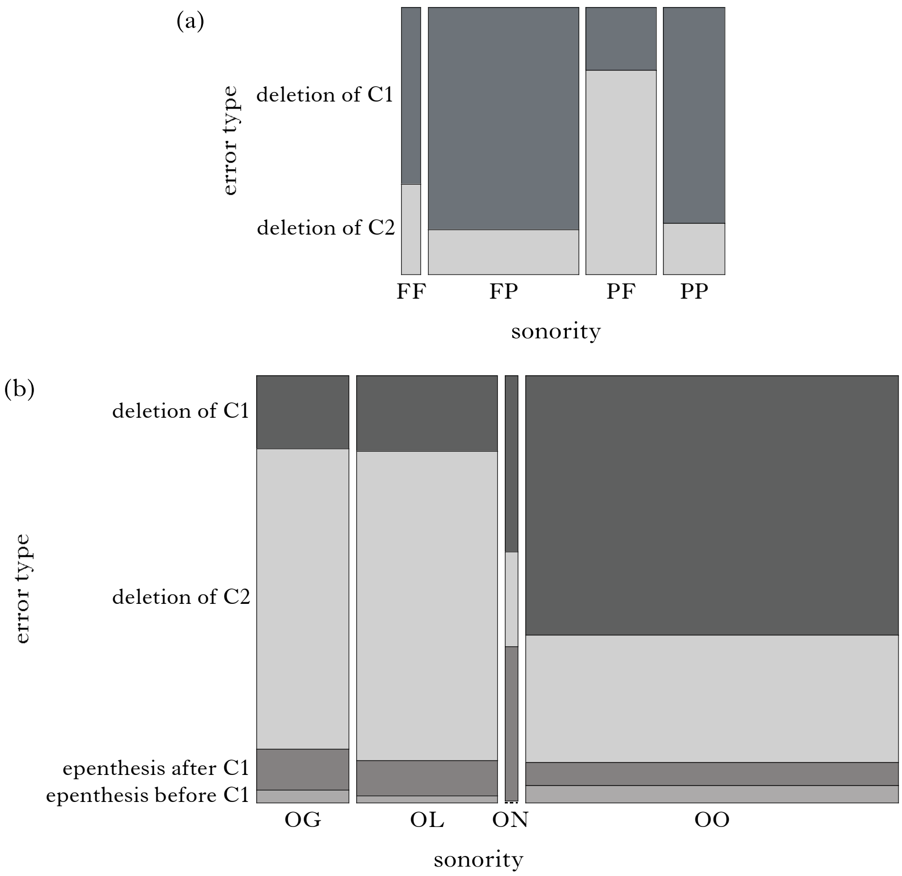

Figure 9
Error types by (a) finer-grained; (b) coarser-grained sonority.

pʂ	1522	zb	88	ʂk	36	xm	17	dl	8	tʂk	3
st	427	zg	86	fs	35	zʐ	17	ʂw	8	tʂʨ	3
kr	411	vr	86	ɕf	34	ftʂ	17	ʣb	8	ʂr	3
sp	370	zr	82	xf	33	dm	16	mn	8	rʐ	2
pj	365	zj	78	vl	33	ʨf	15	tsf	8	gm	2
vj	344	zv	78	xʨ	31	ʑr	15	ʦm	8	rt	2
pr	340	kʲj	75	gv	29	ʂf	15	lv	8	ʑl	2
tr	266	kɕ	74	gl	29	vb	15	gn	7	dn	2
sk	257	bl	73	zl	29	mɲ	15	xʦ	7	tn	2
gr	249	fr	67	ml	29	zʥ	14	tj	7	fʦ	2
mj	248	sx	60	ʣv	28	px	13	tk	7	vʥ	2
sw	227	fp	59	kt	28	ɕp	13	mʂ	7	tʂx	2
br	206	dv	58	pɕ	27	tʂw	12	kn	6	bʑ	2
kl	196	ps	54	pt	27	sʦ	12	bz	6	vm	2
dr	190	bw	52	fʂ	27	zgʲ	12	sʂ	6	mx	2
pw	175	bʐ	51	fɕ	27	gʲj	12	gz	5	ln	2
pl	172	xw	49	vʑ	25	ft	12	wz	5	kʨ	2
vw	146	ʂp	48	fʨ	25	xʂ	11	ʦn	5	lɲ	2
zw	140	zɲ	47	ʂn	24	tx	11	ʐm	4	vɲ	1
bj	133	xr	46	ʂm	23	ɕm	11	stʂ	4	ʐb	1
gw	121	tw	46	dʐ	23	vd	11	db	4	wb	1
kʂ	121	kf	46	ʂt	23	ss	10	tɲ	4	wg	1
zd	119	ɕɲ	46	mr	22	gʥ	10	pɲ	4	zz	1
gʐ	118	ɕr	44	tʂf	22	sɕ	10	fk	4	ʐn	1
zm	117	tf	42	dj	20	ʐw	9	sʨ	4	pʨ	1
ʂtʂ	116	mw	42	sr	20	ʥv	9	xj	3	km	1
zn	107	xl	40	fx	20	tl	9	ʐv	3	mʐ	1
sm	106	dw	40	vn	20	sn	9	ʨm	3	dʐdʐ	1
kw	102	vʐ	39	gd	19	sl	9	rʣ	3	vv	1
tʂ	99	sf	38	ʂl	18	dɲ	8	ʐɲ	3	vg	1
ɕʨ	94	fl	38	tʂt	18	vz	8	rj	3	lj	1
ɕl	89	gɲ	37	fj	18	rv	8	lʐ	3		

Table VIII

Initial segment bigram type frequency in the Polish CDS dictionary.