

Labiodental /r/ project

Experimental design

Background

This project will investigate the acoustic properties of /r/ sounds during their acquisition and test the perceptual abilities of children in relation to different /r/ sounds. Children's production of /r/ and their performance on purposely designed perceptual tests will be recorded at three points during the study. Perceptual tests will also be administered to adults using the same stimuli. This work will gather useful data on the acoustic properties of /r/ in young British children, and will also begin to address why labiodental /r/ users maintain a realization of /r/ which differs from the majority of /r/ tokens in their environment. More generally, the project will investigate the link between production and perception in adults and children.

The details of the experiment including research questions, subjects, testing, recording equipment, and controlled variables are below.

Research questions

The main research questions underlying the study are:

- 1) What are the acoustic properties of /r/ in different structural contexts for British English children aged between 3;0 and 4;0?
- 2) What is the variability (between children and contexts) in the acoustic characteristics of /r/ for children between 3;0 and 4;0?
- 3) What, if any, are the differences in perceptual abilities between children with a labiodental realization of /r/ and children with an apical realization of /r/, and how do those abilities change over the course of acquisition?
- 4) How do the perceptual abilities of children compare to those of adults in relation to /r/-like stimuli?

Subjects

Children

Forty children from nurseries in the South East of England will be tested at three times, each three months apart, between the ages of 3;0 and 4;0. At this stage of development, adult-like /r/ is likely not to have been mastered by the majority of the subjects (Dalston 1975; Strange and Broen 1980; Sanders 1972), although we expect that some children will have acquired adult-like /r/ over the course of the project. The testing of children will consist of a preliminary assessment of articulatory ability, a production elicitation task, and a perception task, details of which are described in the following section.

The subjects will be selected from classrooms in mainstream nurseries and their parents or guardians will be asked to consent to their participation. The number of nurseries has been chosen in order to 1) limit the disruption of classes at any given nursery and 2) control for variation in schooling experience.

It is anticipated that attrition rates and other problems with availability of subjects on an ongoing basis may reduce the number of subjects contributing usable data over the course of the study. For this reason, data will be collected from 40 subjects initially, but the experiment will be viable with half this number.

Adults

Forty adults (20 with labiodental realizations of /r/; 20 with apical realizations of /r/) will be tested once on their perception of /r/-like variants. Subjects will be recruited from the student population of City University.

Test details

Assessment of articulation and phonology

Prior to carrying out the production and perception tests, it is necessary to confirm that all child speakers participating in the study are performing at an age-appropriate level and have no speech disorders other than consistent age-appropriate errors that would bias the data. Therefore, a five-minute ‘Diagnostic Screen’ of articulation, phonology, oro-motor ability, and inconsistency will be administered. This test is the preliminary assessment of the Diagnostic Evaluation of Articulation and Phonology (DEAP; Dodd et al 2002). Predictive ability of the Diagnostic Screen in identifying speech disorders has been confirmed in two independent studies (Dodd et al 2002: 43-44), both of which concluded that the test over-identifies, rather than under-identifies, possible speech disorders. This tendency for the Diagnostic Screen to return a small amount of false positives means that some children possessing no actual speech disorders may be excluded from the experiment, but this is preferable to accidentally testing children with a speech delay.

The Diagnostic Screen consists of a picture-naming task of ten tokens, and guidelines for assessing performance. See Appendix A for the list of tokens and their IPA transcriptions. Two repetitions of the tokens will be recorded and assessed according to the DEAP interpretation guidelines. For the age band of 3;0 to 3;5, subjects are expected to successfully articulate all sounds except / ʃ, tʃ, dʒ, ʒ, θ, ð, r /, and acceptable phonological error patterns include gliding, velar-fronting, deaffrication, cluster reduction, weak syllable deletion, and stopping of fricatives. At the next age band

of 3;5 to 3;11, subjects should have acquired /tʃ/ and ceased the stopping of fricatives.

For both age bands, inconsistency rates should be less than 50%.

Although the DEAP includes additional subtests, the initial Diagnostic Screen is deemed to be appropriate for the purposes of the present study for two reasons. First, it involves fewer test stimuli than the subtests, thus minimizing the time allocated for this portion of the study. Second, as the goal of the present study is not to account for specific error patterns of articulation, phonology, and oro-motor control, the DEAP assessment's role herein is simply that of establishing that subjects lack speech disorders, not of investigating such disorders.

This phase of the experiment will be administered prior to the production and perception tasks on at least 50 children, in order that we may screen for disorders among potential subjects before carrying out the substantive tests that form the core of the present study.

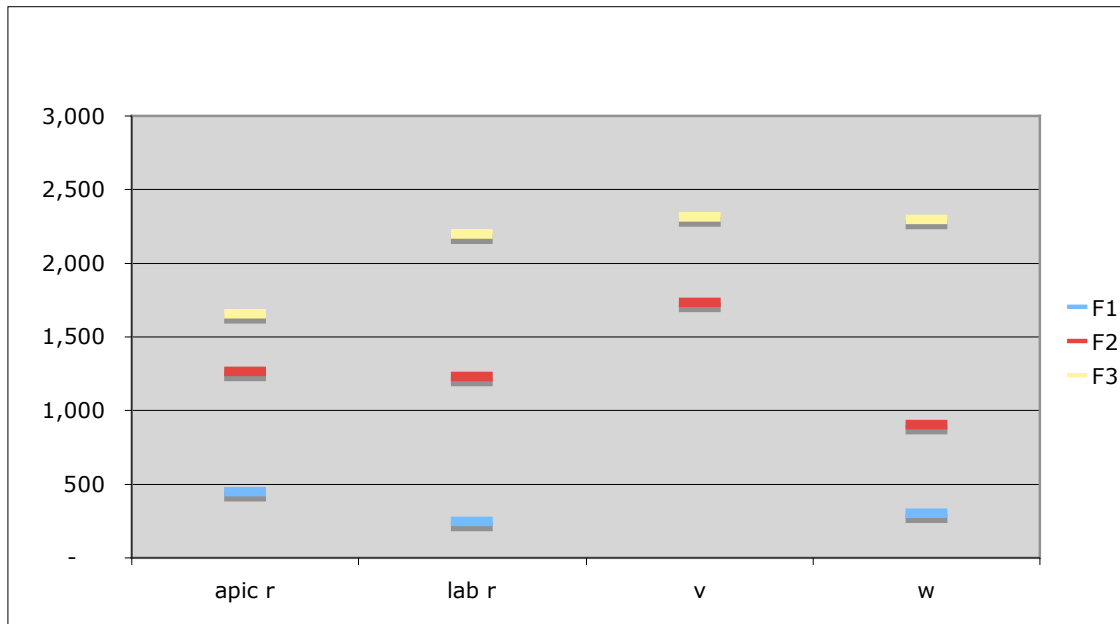
Production task

This task constitutes one of two core experiments. It is concerned with the production of /r/ in three different phonological contexts: word-initially as a singleton; word-initially as the second segment of a consonant cluster; and word-medially as a singleton. In word-initial contexts, /r/ occurs both before the high front vowel /i/ and the low back vowel /a/. The set of tokens also contains two items containing sounds that are similar to /r/ along acoustic, articulatory, and perceptual dimensions: the labial-velar approximant /w/ and the labiodental fricative /v/¹.

¹ It should be mentioned here that wholly fricative and voiced realizations of /v/ may be rare, even in laboratory speech. For many speakers, /v/ tends to be approximantized for at least some of its duration, and

Jones (2005) notes the comparable acoustic data from two speakers (three males and three females with labiodental realizations of /r/). Figure 1 below illustrates average formant values for the male speaker (apical /r/ formant values are those of a different adult male speaker with apical [ɹ], but are included here as a point of comparison).

Figure 1. Acoustic properties of [ɹ], [v], [ʋ], and [w] in /i__i/ contexts (Jones 2005)



Along articulatory dimensions, the four sounds illustrated above overlap in different ways. Labial gestures are present in each of the four, and manner of articulation is consistent for the three approximants in the group. Place features are identical for [v] and for [ʋ].

With respect to perceptual contrasts, it has been noted that the /r/-/w/ contrast is difficult for children to differentiate (Strange & Broen 1980: 138; Dalston 1975) and that misarticulated /r/ and labiodental /r/ are often perceived as /w/ by children (Hoffman et al

in some cases is wholly approximantized. Ohala (1983) also discusses the articulatory difficulty of voiced fricatives due to aerodynamic constraints.

1983) as well as by adults (Dalston 1975). The perceptual link between labiodental /r/ and /w/ is further exemplified by the substitution of the grapheme *w* or the phone [w] for [v] in various literary, cultural, and linguistic sources (see Foulkes and Docherty 2000: 31-32 for a complete account). Although these perceptual generalizations should not be taken as evidence that misarticulated /r/ or labiodental /r/ are identical to /w/, they do indicate that certain /r/ variants may be /w/-like, presumably as a result of similar F3 values. The difference, however, between other formant values of labiodental /r/ and of /w/ (see Figure 1) serves as a clear indication that the two sounds diverge along articulatory and acoustic parameters, and should not be collapsed, regardless of the subtlety in perceptual contrasts.

Given the above-stated similarities among the set of sounds [ɹ, v, w, ɹ], and the questions surrounding the rhotic origin of [v], the inclusion of acoustic data on /w/ and /v/ production is an important aspect of the present study.

Production tokens also include several fillers involving non-rhotic consonants and a range of vowels. The 28 tokens (13 with /r/; 2 with /v/ or /w/; and 13 fillers) were chosen on the basis of their ability to be clearly represented by an image, simplicity, and generally high lexical frequency in written and spoken British English (Leech et al, 2001). The list does not include all possible phonological environments in which /r/ may occur. There are a number of reasons for this. First, the goals of this study do not include a complete analysis of /r/ in all contexts (refer to the research questions above for details on the study's goals). Second, the age of the subjects precludes lengthy testing sessions and it is deemed that the cost of disfluencies, inaccuracies, and gaps in the data outweighs the marginal benefit of assessing /r/ production in every occurring

phonological environment. Third, as is the case with many acoustic studies, token lists must be kept to a manageable number in the interest of time and manpower constraints. Finally, for statistical purposes we have limited the number of independent variables arising from phonological context in favour of increasing the number of repetitions per subject. Appendix B lists the tokens for the production phase of the experiment, along with their lexical frequency counts and IPA transcriptions.

The method of elicitation will be a picture-naming task where images appropriate for 3-4 year-old children are presented in slide-show format on a computer. Various elicitation strategies may be used, but the investigator will avoid pronouncing the token word. A brief warm-up period consisting of informational questions posed by the researcher will precede the experiment so that recording levels may be adjusted and subjects may acclimatize themselves to the presence of recording equipment.

Tokens will be presented in a fixed order for all subjects², with one filler placed between each test token as a distraction. Three repetitions will be elicited, yielding 42 analyzable /r/ tokens for each subject at each of the three testing phases (1,680 tokens at each phase).

Perception task

The second of the two core experiments involves testing the perceptual abilities of children and adults using synthesized tokens in which the formant values of F2 and F3 are precisely manipulated. The stimuli consist of a set of three copy-synthesized nonsense words based on the production of /riba/ and /wiba/ by an adult male with apical [ɹ]. These tokens will be presented to the subjects in XAB format using the Speech Input

² Although we expect each subject to complete the picture-naming task, we have chosen to present the tokens in a fixed sequence to ensure comparability among subjects in the event some do not finish the experiment.

Processing in Children (SIPc) interface (Vance et al, 2005). Formant synthesis was performed at the University of Cambridge using Procsy (Heid and Hawkins, 1998) and HL Syn, allowing the alteration of formants during the steady-state portion of the /r/ and automatic interpolation of formants during the vowel-consonant transitions.

The resulting token list is as follows:

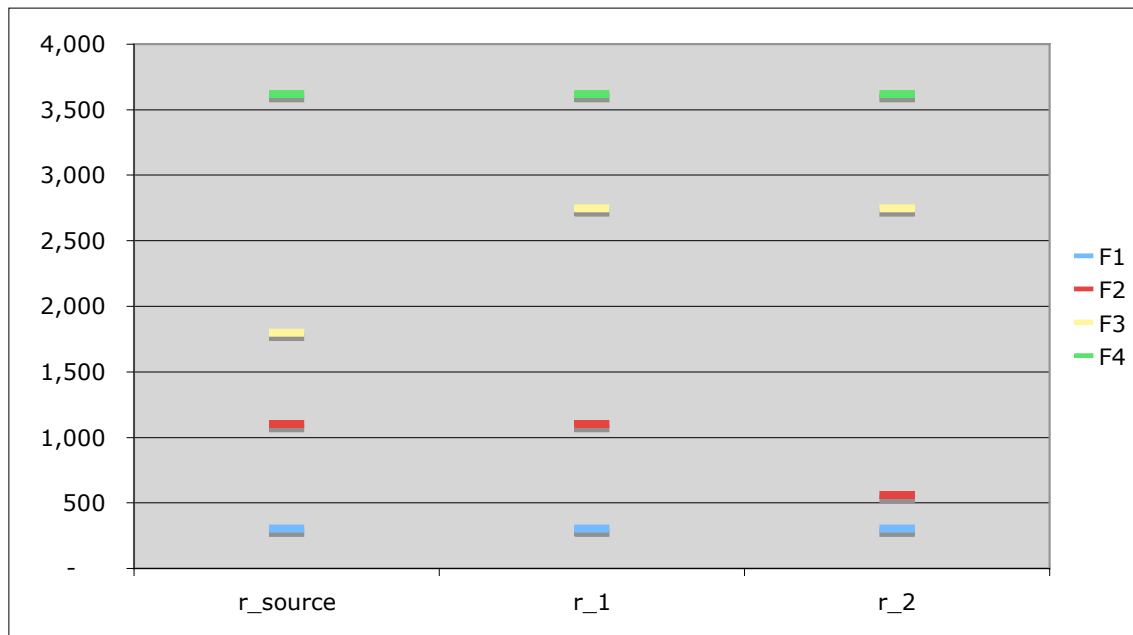
- 1) /r/ with unchanged formants (apical [ɹ])
- 2) /r/ with F3 raised to F3 of /w/ ([ʊ]-like variant)
- 3) /r/ with lowered F2 and raised F3 ([w]-like variant with /r/-like duration)

Table 1 lists the formant frequencies for these synthesized stimuli, and Figure 2 illustrates the relationship among the five stimuli graphically.

Table 1. Formant frequencies for XAB stimuli

Stimulus	F1	F2	F3	F4
r_source	300	1,100	1,800	3,617
r_1	300	1,100	2,743	3,617
r_2	300	557	2,743	3,617

Figure 2. Formant frequencies of XAB stimuli in /__iba/ contexts



These three stimuli were chosen to yield the smallest possible number of XAB combinations (ordered A-B pairs where every stimulus occurs as X four times) that captures the broad categories of occurring sounds [ɹ], [w], and [v]. Using this stimulus set, 12 XAB/XBA combinations are derived. Subjects will be presented with three repetitions of these combinations, for a total of 36 perception tokens per subject at each testing phase (adult subjects will undergo only one phase of testing).

Recording and equipment

DEAP screening and production data will be recorded in quiet rooms familiar to the subjects using a cardioid condenser microphone and a Marantz PMD671 portable digital recorder. Images for the DEAP Diagnostic Screen will be presented using the DEAP stimulus book; images for the production task will be presented in slide-show

format on a laptop computer. These consist of colorful, child-appropriate, and uniformly-sized cartoons collected from various flash-card websites.³

Perception data will be presented and tracked using the SIPc interface on a laptop computer.

Independent Variables

The following independent variables are controlled in the experiment.

- Phonological environment (initial singleton, initial cluster, medial singleton)
- Vowel type (high front, low back – for initial consonant environments only)
- Subject gender
- Subject age
- Subject /r/-style (apical or labiodental)

³ <http://mes-english.com/flashcards> and <http://eslhq.com/gallery>

Appendix A – DEAP Diagnostic Screen Token List

Token	IPA
watch	/wɒtʃ/
fishing	/fɪʃɪŋ/
gloves	/glʌvz/
spider	/spɑɪdə/
thank you	/θæŋkju/
scissors	/sɪzəz/
helicopter	/helɪkɒptə/
bridge	/brɪdʒ/
umbrella	/ʌmbrelə/
elephant	/eləfənt/

Appendix B – Production Experiment Token List

Core experiment tokens

Token	IPA (OED 1989)	Frequency (Leech et al 2001)
branch	/brɑ:ntʃ/	85
breathe	/bri:ð/	35
dry	/draɪ:/	56
dream	/dri:m/	26
grass	/grɑ:s/	45
green	/gri:n/	101
run	/rʌn/	406
rock	/rɒk/	93
read	/ri:d/	284
mirror	/mɪrə/	43
orange	/ɒrɪndʒ/	17
submarine	/sʌbməri:n/	--
kangaroo	/kæŋgəru:/	--
weak	/wi:k/	45
vee (letter)	/vi:/	51

Fillers

cow	duck	moon
nurse	fish	dog
sun	blue	comb
hat	house	king
socks		

References

- Dalston, R.M. 1975. "Acoustic characteristics of English /w, r, l/ spoken correctly by children and adults." *Journal of the Acoustical Society of America*, 57: 462-469.
- Dodd, B., Z. Hua, S. Crosbie, A. Holm, and A. Ozanne. 2002. *Diagnostic Evaluation of Articulation and Phonology*. London: Psychological Corporation.
- Foulkes, P. and G.J. Docherty. 2000a. "Another chapter in the story of /r/: 'labiodental' variants in British English." *Journal of Sociolinguistics*, 4: 30-59.
- Heid, S. and S. Hawkins. 1998. "PROCSY: A hybrid approach to high-quality formant synthesis using HLsyn." In *Proceedings of the 3rd ESCA/COCOSDA International Workshop on Speech Synthesis*, Jenolan Caves, NSW, Australia.
<http://kiri.ling.cam.ac.uk/procsy>
- High Level Parameter Speech Synthesis System (HLsyn). Sensimetrics Corporation.
<http://www.sens.com/hlsyn/index.htm>
- Hoffman, P.R., S. Stager, and R.G. Daniloff. 1983. "Perception and production of misarticulated /r/." *Journal of Speech and Hearing Disorders*, 48: 210-215.
- Jones, M. .2005. "An acoustic study of labiodental /r/ in British English." In F. Chalcraft and E. Sipetzis (Eds.) *Cambridge Occasional Papers in Linguistics*, 3: 100-119.
- Leech, G., P. Rayson, and A. Wilson. 2001. *Word Frequencies in Written and Spoken English: based on the British National Corpus*. London: Longman.
- Ohala, J. 1983. "The origin of sound patterns in vocal tract constraints." In P.F. MacNeilage (Ed.), *The production of speech*. New York: Springer-Verlag. 189-216.
- Sanders, E.K. 1972. "When are speech sounds learned?" *Journal of Speech and Hearing Disorders*, 37: 55-63.
- Strange, W. and P.A. Broen. 1980. "Perception and production of approximant consonants by 3-year-olds: a first study." In G.H. Yeni-Komshian, J.F. Kavanagh, and C.A. Ferguson (Eds.) *Child Phonology*, Volume 2. New York: Academic Press. 117-154.
- Vance, M., S. Rosen, and M. Coleman. 2005. *Speech input processing in children (SIPc): development of an assessment of speech discrimination – an initial report*.
<http://www.ucl.ac.uk/~sslomag/e2.html>