

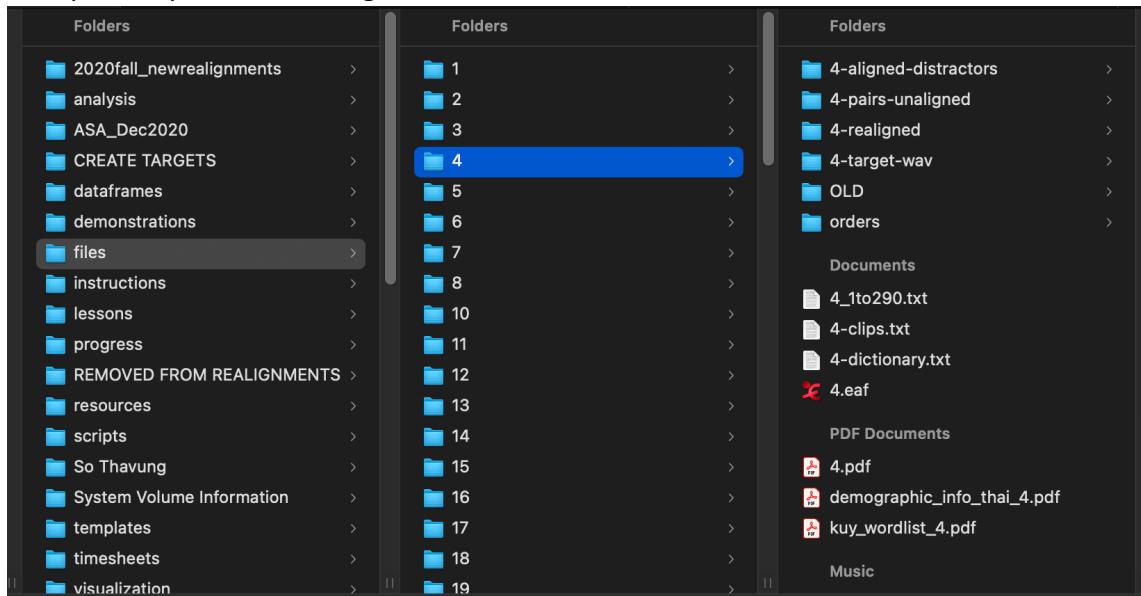
Workflow for gathering and realigning files

Raksit T. Lau-Preechathammarach

2021.01.28

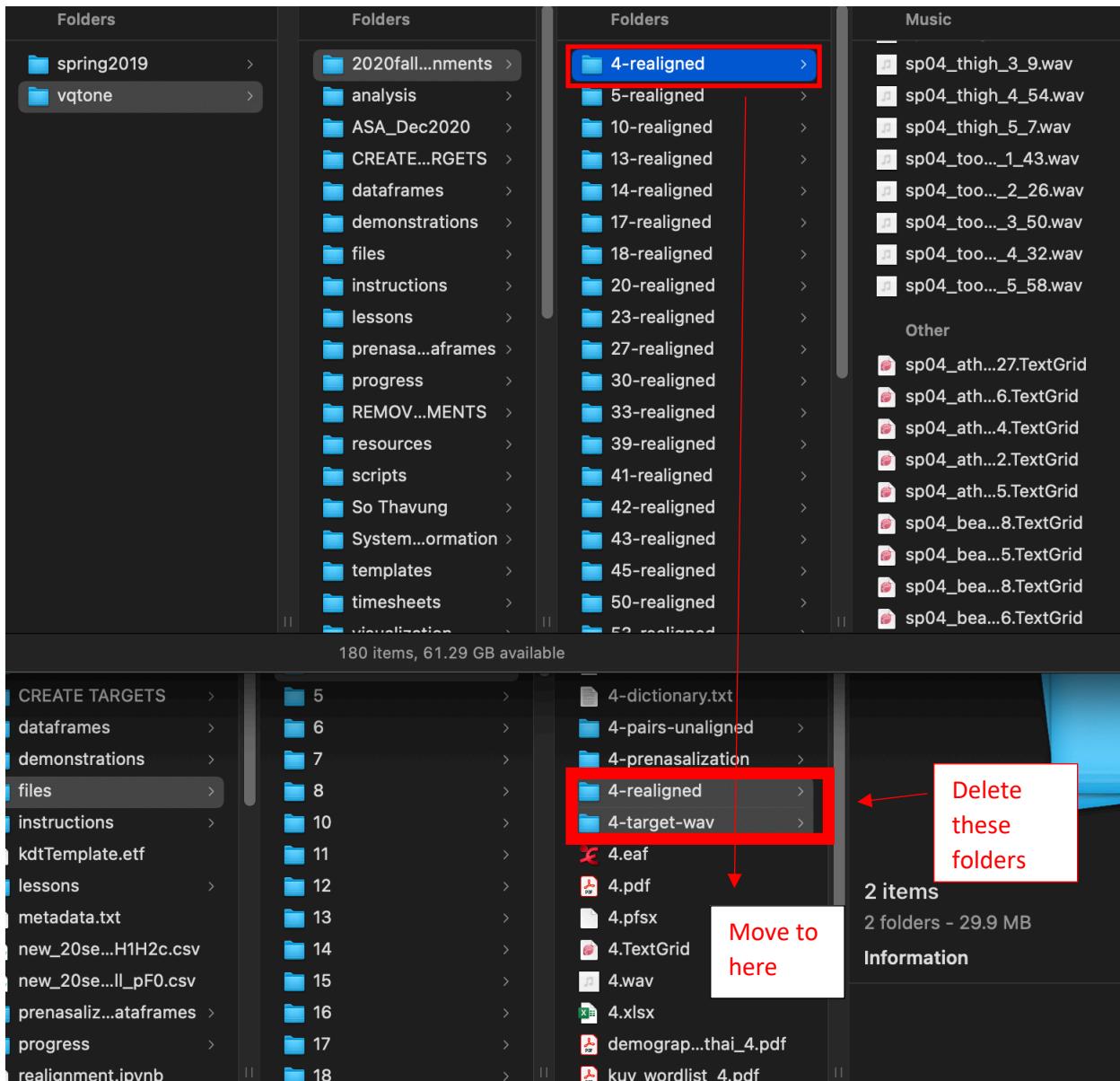
This file documents the steps for gathering and realigning files to look at prenasalization

- (1) You should mainly be working with the speaker folders in **files**. I will be using # to refer to the speaker you are working with



- (2) First let's do a little bit of cleanup

- a. If your speaker was part of the realignment process (i.e. in 2020fall_newrealignments), please
 - i. Delete the **files/#/#-target-wav folder** and **files/#/#-realigned folder**
 - ii. Move the **2020fall_newrealignments folder** to the speaker folder (**files/#**)



- b. If your speaker was **not** part of the realignment process, just go to the speaker folder in **files** and move the .wav files from **#-target-wav** to **#-realigned**. Then delete the target-wav folder

- (3) Now for your speaker, open up the **consolidated** excel file and look at the words that do not have red highlights under the “prenasalized” and “not_prenasalized” columns. These are all the words with potential prenasalization.

meaning	ipa	prenasalized	not_prenasalized	kr	final l	final r	different_word
diligent	cxntrʌŋ	5	0	5	0	0	5
egg	ntræ:l	5	0	5	0	0	0
eggplant	ŋkʌŋ	5	0	5	0	0	0
flower	pi:l	5	0	5	5	0	0
hay	ncha?	2	3	5	0	0	0
hide	nchu:n	0	5	5	0	0	0
just	mpò:m	5	0	5	0	0	0
lice	ncæ:	5	0	5	0	0	0
mom	mpè?	5	0	5	0	0	0
paddy	træ:	5	0	5	0	0	0
redant	ntra:n	5	0	0	5	0	0
sensitive	sxnki:l	4	0	5	4	0	1
shoulder	lxmpa:?	5	0	5	0	0	0
shovel	ntri:m	5	0	5	0	0	0
side	kæ:n	4	1	5	0	0	0
smelly	ncho?	0	5	5	0	0	0
spin (include if pi:l)	5	0	5	5	0	0	0
star	ntɔ:l	5	0	5	0	0	0
stick	rxmpàt	5	0	5	0	0	0
tell	nte:	4	1	5	0	0	0
waist	ŋkæ:n	4	1	5	0	0	0
wrapped	txmpo:m	5	0	5	0	0	0

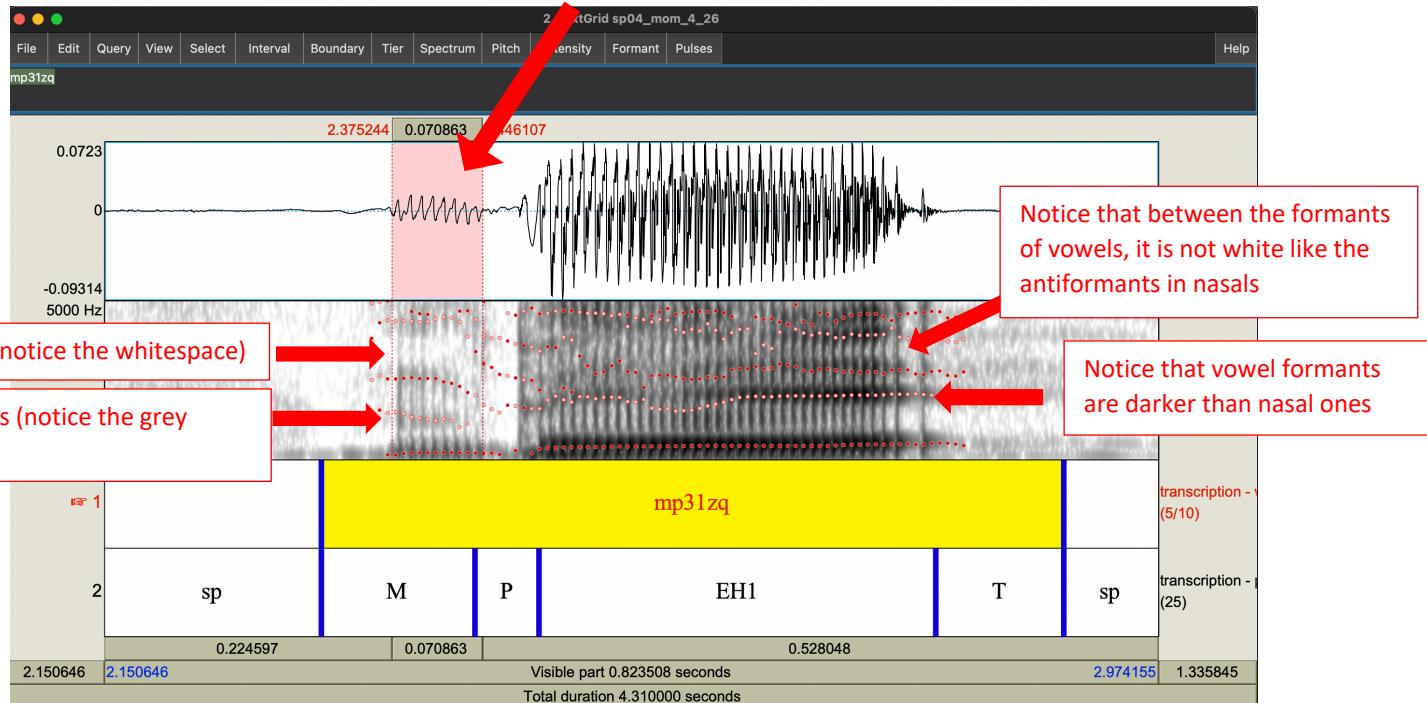
- (4) Now create a folder in the speaker folder labeled #‐prenasalization. Inside it, create a folder labeled #‐prenasalized and one labeled #‐prenasalized-not (where # is the speaker number). Now isolate the .wav and .TextGrid files that correspond to these words. **You should sort isolate these files and put them in some folder first, before moving them into the #‐prenasalized or #‐prenasalized-not folders, so that you don't get lost!**

- If it is a word that was also a target word (this means that it was a target word in the original experiment, which is information you can find in the **stimuli.csv** file), the .wav and .TextGrid files should likely be in the **#‐realigned** folder. If it is, **make a copy of both the .wav and .TextGrid files** and move it into your temporary folder
- It is possible also that the word was a “failed pair”. If this is the case, the files should be in the **#‐failed‐pairs** folder. Some speakers do not have a #‐failed‐pair folder. In these cases, the files should be in **#‐aligned‐pairs**. In this case, **do not make a copy but just move the files** into your temporary folder.
- If the word **was NOT a target word**, the files should be in the **#‐aligned‐distractor** folder. In this case, **do not make a copy but move them** into your temporary folder.

- (5) Now go through your temporary folder and work on realignment. Make sure when you save your work you do not save over the original aligned version.

- If VOT has not already been marked, you should mark it

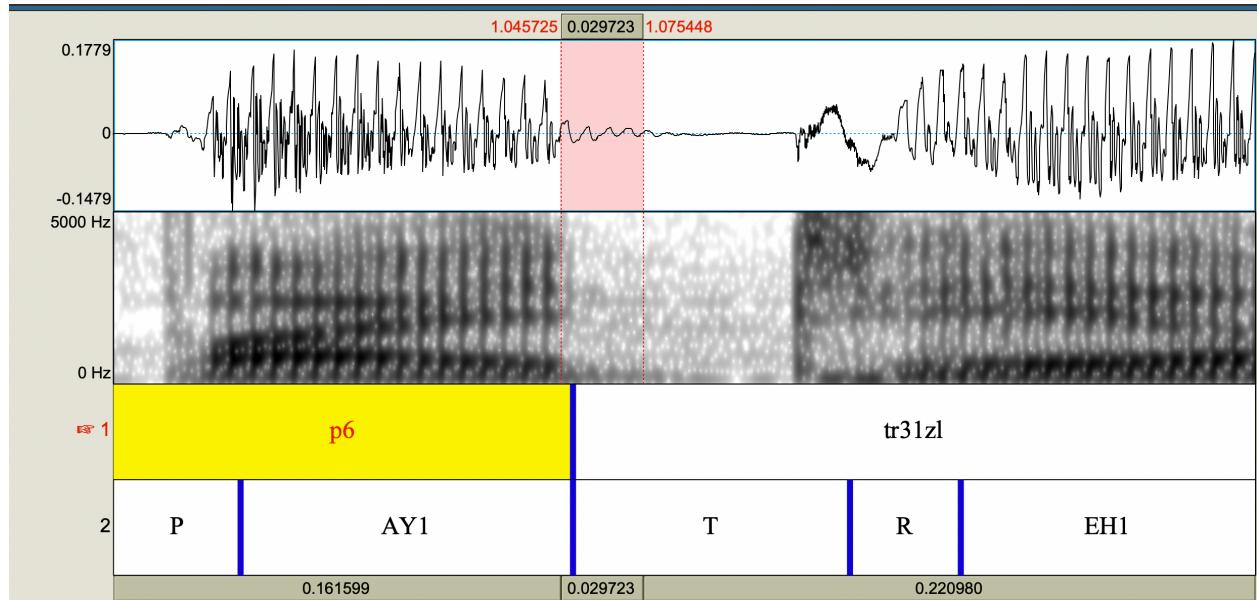
- b. If the vowel has not been realigned yet, you should realign it
- c. If the nasal exists phonetically, you should see a **periodic waveform (voicing), weaker formants than vowels, between formants**



- (6) In the example above, where I would mark the nasal is within the selected interval. Look at the notes to see the cues for the nasals.
- (7) If there is no nasal, you shouldn't see these cues. Sometimes because of the quick transition from the previous vowel, you will see some leftover voicing that might lead you to think it is a nasal (for example, look at the red selection below). But note that there neither formants nor antiformants like the nasal in the above example. Usually you should be able to hear the

Make this ntr31zl →

prenasalization. If you are unsure of anything please please ask me!



- (8) If there is no nasal, you can delete the nasal in the second tier but retain it in the first. Note that in the above example, the nasal is deleted in **both tiers**. This is because early on in my transcriptions, I removed it even from the transcription. Please readd it in the first tier! Here are the transcriptions

IPA	MFA
lu:	luz
lù:	lu1z
ku:	kuz
kù:	ku1z
tpat	tpat
tpàt	tpa1t
lap	l4p
làp làp	l41p l41p
kò:	ko1z
kho:kho:	khozkhoz
pho:m	phozm
mpò:m	mpo1zm
tmo:po:m	tmozm
nte:	ntez
tè:	te1z
the:	thez
po:t	pozt
po:t	po1zt
ti:	tiz

tì:	ti1z
da?	daq
ta?	taq
tà?	ta1q
dàh	da1h
tah	tah
täh	ta1h
kłl	k^1l
khal	khal
to:ŋ	toz9
tò:ŋ	to1z9
pu:?	puzq
pù:?	pu1zq
bu?	buq
pi:l	pizl
pì:l	pi1zl
ncu:n	ncuzn
cù:n	cu1zn
ntræ:l	ntr31zl
mpè?	mpe1q
ncæ:	nc3z
ŋkæ:ŋ	9k3z9
ŋkʌŋ	9k^9
ntri:m	ntri1zm
ntra:ŋ	ntr4z9
rmpàt	rmpa1t
Impa:?	Impazq
ntɔ:l	nt0zl
ktv:	kt8z
sʌ:ŋ	s^z9
nchɔ?	nch0q
ncha:?	nch4zq
cntrʌŋ	cntr^19
phlw:m	phlyzm
sŋki:l	s9kizl
thræ:	thr3z
c̥v:l	c81zl
sæh	s3h
kæ:ŋ	k31z9

- (9) NOTE: Many speakers pronounce the word “mother” as [be:?). If they do, this does not count as prenasalized, but please write in the speaker’s notes which files are pronounced with [b].
- (10) If the token is nasalized, you can move it to the #**-prenasalized folder**. If it is not, you can move it to the #**-prenasalized-not folder**.
- (11) Once you are done, you can **move the original force-aligned files back to their location if they were distractors**. Delete the temporary folder and then put your initials in the **progress_prenasalization.xlsx** file.

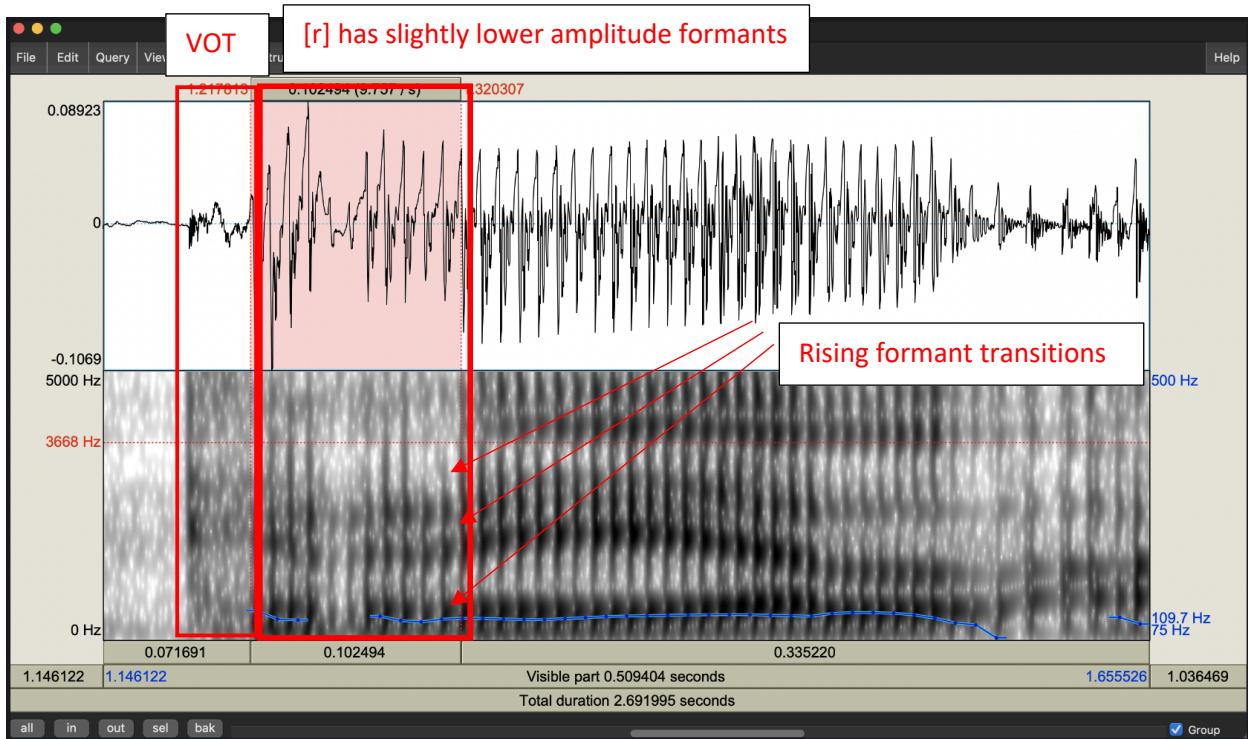
QUESTIONS:

- (1) If a word is in both the realigned folder and another folder, should we realign both?**

If it is in both folders, take the **latest realigned version** (the one that you all worked on, so the vowel should already be realigned and the VOT should be there) and realign from there, but keep the copy that doesn’t have the nasal realigned

- (2) What’s a good way to differentiate the [r] from the vowel (like in the word for ‘egg’)?**

Check out the graphic below. The red area is where I would say the [r] is in [ntrə:l] ‘egg’. You can see that just before is the VOT of [t]. The middle lighter area is part of the trill. The big trick for [r] is that [r] tends to lower all three formants, so you should generally see the formants rising as it transitions into the vowel. You can set the boundary at the end of a cycle where the transition seems to be happening (and when you listen the trill will likely be gone or at the least, faint)



(3) What do we do about words that are nasalized but the sound before it is missing? Should we remove it from the second tier? I guess in general, any missing sound.

Please remove from the second tier! But keep it in the first tier (or restore it if it is gone). For example, if the first tier says ntr31zl but there is no nasal, you can remove the N in the second tier, but keep the first tier as is. If it says tr31zl though, please make it ntr31zl.

(4) If a token is not prenasalized, do we still need to realign?

If it is not, please realign the VOT and the vowel!