Variable phonological phenomena in speech perception:
regressive voicing assimilation and the perception of voicing contrasts in Spanish

Recent studies have argued that speech perception entails an integration of acoustic/auditory cues and knowledge of phonological processes (Durvasula et al. 2018, Cavirani and Hamann 2022). This approach is based on data that shows that phonological processes can determine speech perception despite the presence of acoustic cues that might be in conflict with the perceived sound. Assimilation, in particular, has been found to determine perception of a given sound, even if the acoustic signal corresponds to a different one (Snoeren et al. 2006, Gow & Im 2004, Mitterer et al. 2013, Cavirani and Hamann 2022). However, most studies focus on the perception of the target sound, and consequently, the impact of assimilation on perception of the trigger is not as well understood. In addition, while previous work has shed light on instances where the target sound is contrastive for the assimilated feature, less is known about cases where the alternation is only allophonic (but see Meunier 1999). This study expands our understanding of the role of phonological processes in speech perception by focusing on regressive voicing assimilation (RVA) of /s/ in Spanish and examining whether this process impacts the perception of the voicing contrast of the triggering stop. This process is of interest because it has been shown to be variable and /s/ is not contrastive for voicing in Spanish.

Spanish RVA targets /s/ when followed by a voiced consonant, within and across words (e.g. [las kamas] "the beds" vs. [laz ɣamas] "the ranges") and has been described as variable and gradient (Campos-Astorkiza 2019, Sedó et al. 2020). Among the trigger consonants, stops are contrastive for voicing in Spanish and thus, allow us to examine the role of /s/ voicing in the perception of that contrast. Moreover, we investigate whether the voiced and voiceless allophones show different effects, to explore the impact of variability, and any perceptual differences based on the stop's place of articulation (POA). To answer these research questions, we carried out a forced-choice identification task where participants listened to a combination of two words, with the assimilation context at the word boundary, and had to decide which words they had heard. The stimuli were created using recordings of 10 minimal pairs such as /la**s** bekas/ "the scholarships" vs. /la**s** pekas/ "the freckles" (1) produced by a Castilian Spanish speaker and manipulating the voicing during /s/ using Praat to create 2 combinations where /s/ matched in voicing with the following stop ([s]+voiceless stop and [z]+voiced stop), as expected given RVA, and 2 combinations where /s/ and the following stop didn't match in voicing ([s]+voiced stop and [z]+voiceless stop). There was a total of 40 test items, plus 40 fillers. Data was collected via Qualtrics from 120 participants from central and northern Spain. Each data point was coded for accuracy according to the voicing of the stop, and we examined the impact on accuracy of the stimuli structure ([s]+voiceless, [s]+voiced, [z]+voiceless, or [z]+voiced), the stop's POA, and their interaction with logistic regression in R.

Overall, results show that stimuli with combinations mismatched for voicing present the lowest accuracy rate (Table 1). For those combinations, listeners can perceive the stop according to the voicing during /s/ rather than the stop's voicing, indicating that in those cases, listeners are using their knowledge of RVA to perceive the stop instead of using its voicing cues. However, results show an asymmetry for the mismatched stimuli: for [s]+voiced, listeners show a much higher percentage of correct responses than for [z]+voiced. For this combination, listeners perceived a voiced stop even though its acoustic signal was voiceless at a rate of 31%. This asymmetrical pattern can be interpreted as stemming from the variable nature of RVA in Spanish. In production, the voiceless allophone [s] may be followed by a voiceless stop or a voiced one, when RVA is not applied. On the other hand, the voiced allophone [z] only occurs when the following consonant is voiced (Campos-Astorkiza 2019). This variability is reflected on the perception patterns: [z] may lead to perception of a voiced stop, despite the presence of cues for its voicelessness, in almost a third of the data. The stop's POA also impacts the accuracy rate (Table 1). The dental POA displays the lowest accuracy

rate and the greatest asymmetry for the stimuli mismatched for voicing. These place-based differences might stem from articulatory-acoustic differences among the three stops since the Spanish voiced dental stop usually presents less constriction compared to the other POAs (Colantoni & Marinescu 2010), which might result is more auditory cues for its voicing.

To conclude, this project brings new evidence on the role of phonological knowledge on speech perception by focusing on RVA and builds off previous studies by analyzing variable RVA and its impact on the perception of the trigger consonant.

(1) Test items

/las batas/ ~ /las patas/	'the robes, the legs'
/las barkas/ ~ /las parkas/	'the boats, the parkas'
/las bekas/ ~ /las pekas/	'the scholarships, the freckles'
/los bo̞jos/ ~ /los po̞jos/	'the pastries, the chickens'
/las dunas/ ~ /las tunas/	'the dunes, the music groups'
/las domas/ ~ /las tomas/	'you tame them, you take them'
/las komas/ ~ /las gomas/	'you eat them, the erasers'
/las kotas/ ~ /las gotas/	'the levels, the drops'
/las kalas/ ~ /las galas/	'the coves, the galas'
/las kasas/ ~ /las gasas/	'the houses, the gauzes'

Table 1. % of correct responses for all data (*Total*) and by stop place of articulation according to the structure of the stimuli

	Stimuli structure	**Total**	**b/p**	**d/t**	**g/k**
Matched for voicing	[s]+voiceless stop	92.83%	91.25%	90.83%	95.42%
	[z]+voiced stop	97.83%	99.58%	92.5%	98.75%
Mismatched for voicing	[s]+voiced stop	90.33%	88.75%	95%	89.58%
	[z]+voiceless stop	69.00%	81.25%	56.66%	62.92%

References

Campos-Astorkiza, R. (2019) Modelling assimilation: The case of sibilant voicing in Spanish. In J. Gil & M. Gibson (eds.) *Contemp. Studies in Romance Phonetics & Phonology.* 241-275.

Carivani, E. & S. Hamman (2022) Formalising phonological perception: The role of voicing assimilation in consonant cluster perception in Emilian dialects. *J. Linguistics*, 1–32.

Durvasula, K. et al. 2018. Phonology modulates the illusory vowels in perceptual illusions: Evidence from Mandarin and English. *Laboratory Phonology 9*

Gaskell, M. G. (2003). Modeling regressive and progressive effects of assimilation in speech perception. *J. Phonetics* 31, 447–463.

Gow, D. W. & A.M. Im (2004). A cross-linguistic examination of assimilation context effects. *J. Memory and Language* 51, 279–296.

Mitterer, H., S. Kim & T. Cho. (2013) Compensation for complete assimilation in speech perception: The case of Korean labial-to-velar assimilation. *J.Memory & Lang.* 69, 59–83.

Snoeren, N. D., P.A. Hallé and J. Segui. (2006). A voice for the voiceless: Production and perception of assimilated stops in French. *J. Phonetics* 34, 241–268.

Meunier, C. (1999) Recovering missing phonetic information form allophonic variation. In *ICphS 1999*.

Sedó, B., L. Schmidt and E. Willis. (2020) Rethinking the phonological process of /s/ voicing assimilation in Spanish: An acoustic comparison of three regional varieties. *Studies in Hispanic and Lusophone Linguistics* 13, 167–217.