Sound change by low surprisal: the case of Portuguese /ẽw̃/

One of the most controversial and yet-to-be-resolved issues in the historical phonology of Portuguese refers to the emergence of its word-final nasal diphthong /ẽw̃/, spelled -ão (e.g. mão [mẽw̃] 'hand') or -am (e.g. amam [ˈɐmẽw̃] 'they love') in the modern language. The intense debate over this issue has centered on the two phonological mergers that gave rise to this diphthong. Until the 1300s, it is hypothesized that Old Portuguese had the following three distinct word-final nasal sequences: /ẽᵑo/, /ẽᵑ/, and /õᵑ/, e.g. mão [ˈmẽᵑo] 'hand', pam [ˈpẽᵑ] 'bread', razom [rɐˈdzõᵑ] 'reason'. By the early 1400s, however, these three sequences had developed a labial-velar offglide, which led the first two sequences to merge in /ẽw̃/, i.e. /ẽᵑ/, /ẽᵑo/ > /ẽw̃/, but /õᵑ/ > /õw̃/. Finally, by the early 1500s, words with final /õw̃/ had joined those ending in /ẽw̃/, thus completing the merger of the three original nasal sequences and rendering modern forms such as mão, pão, and razão, all pronounced with final [ẽw̃] in the standard variety of the language. The origins of /ẽw̃/ in Portuguese have received much philological scrutiny since at least the end of the 1800s (cf. Meyer-Lübke 1890). Throughout the 20th century, however, multiple scholars have attempted to provide an explanation for the aforementioned mergers. To date, proposals have followed mainly two perspectives, namely, analogical (e.g. Williams 1933, 1962; Tilander 1959) and phonetic/phonological (e.g. Nobiling 1903; Sampson 1983, 1999; Carvalho 1989; Fagan 1992). While the former provide implausible diachronic pathways given the available evidence for the changes in question, studies which adopt the latter perspective have at best only described the mergers through ad hoc lists of rules of glide insertion and vowel dissimilation, thus failing to provide a principled explanation as to *why* such rules would have come about in the first place. While sound change has been studied from several perspectives throughout the years, a fundamental principle of language as a communication system has often been overlooked, namely, the information conveyed by language sounds in a given message during spoken communication. Thus, an increasing number of studies have focused on exploring the role that information plays in the shaping of sound patterns (Hume & Mailhot 2013; Cohen Priva 2017; Cohen Priva & Jaeger 2018; Hall et al 2018; etc.). Specifically, the theoretical tools of probability, entropy, and surprisal, as defined under the tenets of Information Theory (Shannon 1948), are crucial to understanding the motivation behind phonological changes, as they shed light upon the nature of the change targeting unstable language patterns (Hume & Mailhot 2013: 44). The goals of the present study are, thus, twofold: (i) to present a unified account of the changes that led to the emergence of final /ẽw̃/ in the history of Portuguese, consistent with a quantifiable theory of phonological change; and (ii) to apply information-theoretic tools as a contribution to the advancement of theoretical approaches to sound change. By analyzing data from a corpus of historical Portuguese (Davies & Ferreira 2006-), this study proposes that the change events incurred by the Old Portuguese word-final nasal sequences derive not only from a likely phonetic ambiguity that characterized each nasal subsystem in the different historical periods in question, but also—and decisively—from the instability of high frequency sounds in those subsystems. In information-theoretic terms, sounds with low surprisal (or high expectedness) are characterized by high frequency, simple articulations, and weak perceptual distinctiveness, and tend to be unstable on many fronts. In particular, changes in sound patterns with low surprisal are usually prone to production-based processes and need not be structure-preserving. This is the proposed motivation for the diphthongization in the changes /ẽᵑ/ > /ẽw̃/ and /õᵑ/ > /õw̃/ in the 14th century, as well as for the dissimilation in /õw̃/ > /ẽw̃/ during the 15th century, which led to the unification of all original nasal sequences in /ẽw̃/ in the standard variety of Portuguese at the beginning of the 16th century.

References

Carvalho, Joaquim Brandão de. 1989. L'origine de la terminaison *-ão* du portugais: une approche phonétique nouvelle du problème. *Zeitschrift für Romanische Philologie* 105. 148-160.

Cohen Priva, Uriel. 2017. Informativity and the actuation of lenition. *Language* 93(3).569-597.

Cohen Priva, Uriel and T. Florian Jaeger. 2018. The interdependence of frequency, predictability, and informativity in the segmental domain. The role of predictability in shaping phonological patterns. *Linguistics Vanguard* 4(s2).

Davies, Mark and Michael Ferreira. 2006-. *Corpus do Português: Historical Genres*. Available online at http://www.corpusdoportugues.org/hist-gen/.

Fagan, David S. 1992. On the history of Portuguese /ẽu, ẽi/: The case for spontaneous diphthongization. *Word* 43(3).383-397.

Hall, Kathleen Currie, Elizabeth Hume, T. Florian Jaeger, and Andrew Wedel. 2018. The role of predictability in shaping phonological patterns. *Linguistics Vanguard* 4(s2).

Hume, Elizabeth & Frédéric Mailhot. 2013. The role of entropy and surprisal in phonologization and language change. In Alan C. L. Yu (ed.), *Origins of sound patterns: Approaches to phonologization*, 29-48. Oxford: Oxford University Press.

Meyer-Lübke, Wilhelm. 1890. *Grammaire des langues romanes*, Vol. I: Phonétique. Paris: Welter.

Nobiling, Oscar. 1903. Die Nasalvokale im Portugiesischen. *Die Neueren Sprachen* 11.129-153.

Sampson, Rodney. 1983. The origin of Portuguese *-ão*. *Zeitschrift für Romanische Philologie* 99.33-68.

Sampson, Rodney. 1999. *Nasal vowel evolution in Romance.* Oxford: Oxford University Press.

Shannon, Claude. 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27.379-423, 623-656.

Tilander, Gunnar. 1959. Porque *-am*, *-om* se tornaram *-ão* em português? *Revista de Portugal*, *Série A: Língua Portuguesa* 24.292-303.

Williams, Edwin B. 1933. The Portuguese final *-ão*. *Language* 9.202-206.

Williams, Edwin B. 1962. *From Latin to Portuguese: Historical phonology and morphology of the Portuguese language*. Philadelphia: University of Pennsylvania Press.