《外國語文研究》第四十一期
2025 年 6 月頁 35-65
DOI:10.30404/FLS.202506_(41).0003

語境與語言限制對擬聲詞「Click」和「Knock」的語音與知覺影響

Kevin Samejon*

波士頓大學

摘要

儘管擬聲詞在不同語言中普遍存在，其本質仍未得到充分認識。這可能是由於對同一聲音的表現或模仿程度不同，結果往往顯得隨機。然而，跨語言研究證據顯示，在特定語境中，某些音段在表現聲音象徵詞義方面具有一定的一致性。基於這些觀察，本研究探討了宿霧語中「Click」和「Knock」擬聲詞的生成與知覺（n=38 位宿霧語母語者）。通過宿霧語的聲學語音分析以及宿霧語和西班牙語的知覺測試，我們的結果顯示跨語言和語言特定的限制結合起來，影響了宿霧語使用者對「Click」和「Knock」擬聲詞的發音與知覺。特別是，擬聲詞的形成和常規化可能促成了受觀察的語言形成特定的模式。

關鍵詞：擬聲詞、宿霧語、西班牙語、言語生成與知覺

* 波士頓大學語言學系博士畢業生

Phonetic and Perceptual Consequences of Context and Linguistic Constraints in *Click* and *Knock* Onomatopoeias

Kevin Samejon*

Boston University

Abstract**

The nature of onomatopoeias is still not adequately understood, despite their crosslinguistic presence. This may be due to the varying degrees of representations or imitations of the same sound, which often appear arbitrary. However, crosslinguistic evidence suggests that certain sound segments exhibit some consistency in representing sound symbolic word meanings within particular contexts. With these observations in mind, the present study explored the production and perception of click and knock onomatopoeias in Cebuano (n=38 native Cebuano speakers). Using acoustic phonetic analysis in Cebuano and perceptual tests in Cebuano and Spanish, our results suggest a combination of crosslinguistic and language-specific constraints that influence Cebuano speakers' production and perception of click and knock onomatopoeias. In particular, onomatopoeic word formation and conventionalization may have contributed to the language-specific patterns observed.

Keywords: onomatopoeia, Cebuano, Spanish, speech production and perception

---

* PhD Graduate of Department of Linguistics, Boston University

** Thanks to the audience of the Sound Symbolism Workshop at the 56th Annual Meeting of the Societas Linguistica Europaea (SLE) for their insightful comments on an earlier version of this paper.

1. Introduction

Sound symbolism refers to the less arbitrary relationship of sound and meaning in human languages, and one class of words that exemplifies this notion is onomatopoeia (cf. Saussure 1959). Onomatopoeic words are sound symbolic imitations that map the acoustics of extralinguistic sound sources to the spoken human language. Being part of human language, onomatopoeias exhibit language-specific constraints, especially along the lines of their phonemic structure (Nuckolls 1999; Güldemann 2008; Körtvélyessy 2020), but the same structure on the phonetic level is hardly ever considered (but see Assaneo et al. 2011). Building on our understanding of the imitative nature of onomatopoeias and observed language constraints, we examine how different imitation contexts and linguistic factors influence the production and perception of the crosslinguistically attested click and knock onomatopoeias.

Our starting point is the widely observed phenomenon in sound symbolism where the production of certain extralinguistic sounds are associated to specific contexts, e.g., size, shape, texture, intensity, or category (Sapir 1929; Köhler 1970; Perlman et al. 2015) The production of the extralinguistic sounds may also be gradient relative to the degree by which they fit these associations (Shinohara & Kawahara 2010; Kawahara & Braver 2014). Along these lines, we hypothesize that the production of onomatopoeias will also vary based on the given intensity or size of the sound source they imitate, i.e., imitation context, and will be expressed via fine-grained phonetic changes. Not only that these fine-grained phonetic changes will likely conform to the relevant imitation contexts, but we also hypothesize that fine-grained phonetic changes will follow certain linguistic constraints found in the language in question (cf. Braver et al. 2016). In terms of perception, we hypothesize that listeners will be able to associate changes in the phonetic segments they hear to specific sound sources of relevant onomatopoeias crosslinguistically.

The rest of the paper is structured as follows. Section 2 situates our current investigation in the contextual and linguistic parameters through which onomatopoeias are expressed in human language and followed by the structure of click and knock. Moreover, Section 2 contains description of the present study and the research questions that drive our experimental efforts. Section 3 and Section 4 describe the methods and report the results of two experiments in this study, i.e., a production task.

(Experiment 1) and a listening-perception task (Experiment 2), respectively. Section 5 presents the general discussion of the studies' findings. Section 5 also revisits the role of context and linguistic constraints in the production and perception of onomatopoeias. Finally, Section 6 provides the conclusion of the present study.

2. Background

2.1. Context and Constraints

The production of onomatopoeias fundamentally involves imitating extralinguistic sounds into a communicable form within human language. What this means is that the imitation is "reshaped to fit the sound pattern of the imitator's language" and the preciseness of linguistic imitation is a function of the sound system of a language (Hocket 267). Linguistically, the imitation of the acoustic qualities of extralinguistic sounds is translated into combinations of phonemes within the language, facilitating a close correspondence between sound and meaning (Childs 2015; Johansson et al. 2020; Kwon & Round 2015). Onomatopoeias are categorized under the broader term of sound symbolism due to this inherent correlation between sound and meaning.

A recent volume by Körtvélyessy & Štekauer (2024) has shown that phonemes in onomatopoeias carry a degree of meaning with crosslinguistic patterns. Certain phonemes are more commonly used across languages depending on the supposed sound source they represent. For instance, (voiced) plosive onsets are often associated with sounds like drums, bells, and thunder, while fricative onsets encompass sounds such as wind, snakes, and laughter. Moreover, /u/ is associated with sounds like splash, storm/thunder, and heartbeat, while /i/ and /u/ alternate for sounds like that of a fly or mosquito, with /i/ more frequently associated with sounds like snakes and chiming bells. In contrast to onsets, codas in many onomatopoeias are typically voiceless or nasal sounds.

The function of phonemice representationss, as demonstrated by Körtvélyessy & Štekauer (2024), extends to expressing specific sensory meanings (Jespersen 1933; Ramachandran & Hubbard 2001; Winter et al. 2017). For instance, across many languages, diminutive expressions conveying concepts like physical smallness or quickness often feature high.

front vowels, whereas expressions denoting largeness and related concepts tend to use low back vowels (Ohala 1994; Nuckolls 1999; Haynie et al. 2014). The use of a particular phoneme presupposes that it is part of the language's phoneme inventory (Lavoie 2002).

The association between phonemes and the imitated sounds appears to still adhere to linguistic constraints. The representation of phones that make up the syllable structure of imitated sounds are optimized to approximate extralinguistic sounds. These linguistic constraints may stem from the phonotactics of a language, which refers to the preferred order or combinations of phonemes. For example, in English, *blip* is common and preferred over *lbip* (Berent 2008), but it will be less common and preferred in Japanese to have such consonant clusters. Moreover, Northern Amis, an Austronesian language in Taiwan (Bril 2024), allows the velar nasal /ŋ/ word-initially and word-finally *ngang* "sound of a sudden noise" but only word-finally in English, bang vis-à-vis *ngba*, reflecting the constraints in preferred phoneme sequences across languages.

Körtvélyessy (2020) exemplified this further through comparisons of dog barks: in English, *arf-arf*, French *ouaoua*, and Slovak *hav-hav*. Each language uses different phoneme combinations to represent the sound of a dog bark, demonstrating that certain combinations may better capture the essence of a sound in one language compared to another. Thus, some combinations will vary across languages due to these linguistic preferences and constraints.

The production of onomatopoeias on the phonetic level, however, is not widely explored although context and linguistic constraints seem to affect the realization of onomatopoeias. Gleaning from the phonetic literature, we can expect crosslinguistic tendencies on how languages behave in terms of their phonetic perturbation, e.g., the degree to which vowel quality and duration are changed, especially in expressive contexts. Take the changes in phonetic duration in emphatic contexts as an example. While there is a crosslinguistic preference for a binary short-long duration distinction regardless of context, speakers may manipulate the phonetic duration of sound segments when communicating various levels of emphasis, e.g., "It was *so/soo/sooo/soooo/sooooo/soooooo* creepy". Speakers of English and Japanese diverged from this crosslinguistic preference by exhibiting a three-way and a six-way duration distinction, respectively, relative to various emphasis levels (Braver et al. 2016;

Kawahara & Braver 2014). In contrast, the crosslinguistic preference for a binary short-long distinction persisted in Cebuano (Austronesian > Philippines; Samejon 2019).

Compared to the production of onomatopoeias, our understanding of their perception is even more limited. Notably, the work of Assaneo et al. (2011) stands out in this regard, where they examined the acoustic characteristics of real world sounds of clicking and knocking and compared them to the production of *click* and *knock* onomatopoeias in Spanish. In their study, they isolated the most identifying sound segments of the *click* and *knock* onomatopoeias in Spanish, which they found to be the coarticulated [ik] and [ok] segments, respectively.

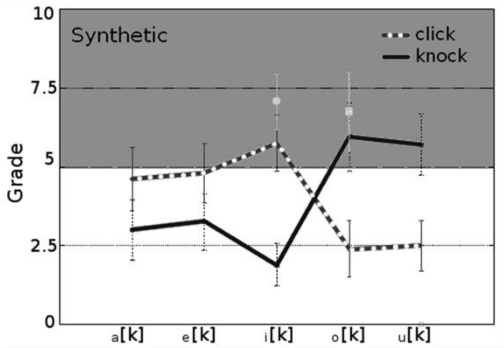

Figure 1. Results from Assaneo et al. (2011) which demonstrated native Spanish speakers gave higher 'grade' or had responded with stronger association on coarticulated [ik] segment for click, and coarticulated [ok] coarticulated segment for knock.

When participants were asked to associate the sounds they heard with either the sound of clicking or knocking as they know them from the real world, native Spanish speakers demonstrated a strong perceptual association of [ik] with clicking and [ok] with knocking, based on the overall mean of their responses. This underscores that onomatopoeias are indeed imitative, and speakers can perceive the key phonetic structures of onomatopoeias (such as the coarticulated segments of [ik] and [ok]) in relation to the imitated real world sound source.

However, since these results were based only on Spanish speakers, their crosslinguistic generalizability remains unclear, despite their discovery that the acoustic composition of the final /k/ and /x/ in click and knock onomatopoeias are analogous to each other. Focusing on the same onomatopoeias as Assaneo et al. (2011), it is natural to ask whether the key identifying sound segments of click or knock in Spanish may be perceived differently by speakers of other languages.

2.2. *Click* and *Knock* Onomatopoeias

The crosslinguistic presence and structural resemblance of click and knock onomatopoeias make them viable candidates for analyzing the phonetic and perceptual realities of onomatopoeic words. These onomatopoeias imitate sounds characterized by short bursts, such as those produced by light switches, computer mouse clicks, and knocking on a door. Furthermore, the communicative use and association of these onomatopoeias with their sound sources have been established in the English language for centuries, dating back approximately 4 to 6 centuries ago (Oxford English Dictionary). Table 1 lists some crosslinguistic onomatopoeic sounds for knocking and clicking. [1]

Table 1. Crosslinguistic examples of the sounds of clicking and knocking.

Language	Sound of clicking	Sound of knocking
Albanian	tik	tok
Bulgarian	trak	chuk
Dutch	tik	klop
English	klik	nok
Hebrew	qliq	tuq
Irish	clic	duñ
Malay	tek	tok
Persian	kilik	taq
Polish	klik	puk

[1] From Assaneo et al. (2011) and Körtvélyessy & Štekauer (2024).

The most common approach to analyzing the data in Table 1 is to identify shared phonetic combinations of *click* and *knock*. Both click and knock sounds feature /k/, which typically appears word-finally but can also occur word-initially. Further distinction between *click* and *knock* comes from the vowels, with [i] predominantly used in linguistic imitations of clicking sounds and [o] for knocking sounds. The short and abrupt acoustic signal of clicking or knocking is effectively captured through the articulation of [k], while the variation in vowels reflects the articulatory imitation of the acoustic resonance of the materials involved (Johansson et al. 2020). Meanwhile, the final coarticulatory segments of these onomatopoeias, [ik] and [ok] are clearly associated with one onomatopoeia over the other. Overall, analysis presented here provides only a helpful and comprehensive description of the phonemic structure of *click* and *knock* onomatopoeias across languages.

Cebuano *click* and *knock* onomatopoeias, *hagtik* and *hagtuk*, respectively, resemble the common crosslinguistic structure of these two imitated sound sources (Wolff 287).[3] These onomatopoeias in Cebuano follow a similar pattern as in other languages. However, they are distinguished by the /i/ and /u/ vowel phonemes in their final syllables. The only alternative, *hagtak*, with the /a/ vowel, represents a cracking or dropping sound. This phonemic contrast may have constrained *click* and *knock* onomatopoeias in Cebuano to have /i/ and /u/, which aligns with the crosslinguistic preference for high front vs. non-high back vowels in imitations of *click* and *knock* sounds across languages. Moreover, the final consonant [k] in both Cebuano *click* and *knock* is unreleased, making this consonant unlikely to be lengthened, but the preceding vowel may likely be lengthened.

The exercise above suggests that a phonemic and segmental analysis of onomatopoeias may be quickly exhausted. While the phonetic consequences of the production of *click* and *knock* may be assumed, additional subsegmental investigation will be necessary. Such pursuit will help pinpoint how the imitation of extralinguistic sounds operates specifically at the acoustic level. When considering perception, we may hypothesize that the high front vowel /i/ is imitative of clicking sounds,

[3] Wolff (1972) defined these words as: *hagtik*, sharp clicking sound; make a clicking, ticking sound, and *hagtuk*, loud knocking sound; produce a loud knocking sound.

while the non-low back vowels /o/ and /u/ function similarly for knocking sounds. If the question were merely about associating which phoneme or sound combinations belong to which real world sound source, the study by Assaneo et al. (2011) has provided answers. However, the question here is broader: whether identifying coarticulated segments like [ik] and [ok] are crosslinguistically perceptible as sounds associated with their respective onomatopoeias.

2.3. Present Study

The study examines the role of context and linguistic constraints in the production and perception of *hagtik* 'click' and *hagtuk* 'knock' onomatopoeias in Cebuano.[4] Despite their imitative nature, we still lack understanding of how the phonetic realization and association of onomatopoeias change relative to the contexts in which speakers attempt to imitate them. Additionally, given that onomatopoeias function within the language system, it is highly probable that both the production and perception of onomatopoeias are influenced by linguistic constraints, whether due to structural reasons or perceptual processing limitations. We provided two experiments to this effect. They are discussed separately in the following sections.

Two experiments are involved in the study. Commonly used phonetic analysis and methods found in the literature (Johnson 2010; Ladefoged & Maddieson 1996) were employed to set up each experiment. Experiment 1 quantified phonetic changes in the onomatopoeias' most distinguishing segment, i.e., the vowel, by considering vowel quality (measured via F1 and F2), pitch (measured via F0), and length (measured via duration), via production task. Experiment 2 examined the perception (or the degree of association between the sound source and the onomatopoeia) of onomatopeias via listening-perception task with an accompanying rating scale. When collecting the data, the experiments described in this paper

[4] An anonymous reviewer suggested the phonetic transcriptions [hag.tak̚] and [hag.tʊk̚] for *hagtak* and *hagtuk*, respectively. While it is not the central occupation of the article, we note that *hagtuk* transcribed as [hag.tʊk̚] may be the potential transcription of its actual production following the claim of Tanangkingsing (2009). However, this vowel change might simply be a case of back vowel lowering as well (Intlekofer & Bishop 2016; Samejon 2022; e.g., Zuraw 2006). To simplify our crosslinguistic comparison, we opted to use [uk] as an approximate production of the back vowel in the final syllable of the Cebuano 'knock', which helped better explain our predictions and results for Experiment 2 later.

were presented to the participants, and upon their interest, they completed a consent form with a clause allowing them to withdraw their participation at any time. All identifying information was replaced with participation IDs, whose codes are only accessible to the researcher, with the code key stored on a secured computer. The participants were duly appreciated for their time and participation, with their contribution highlighted as essential to better understand Cebuano onomatopoeias.

Finally, choosing of the Cebuano language was based on two reasons: accessibility to speakers and its status as an understudied language relative to onomatopoeias. While Cebuano is widely spoken as a native language by a large population in the Philippines (~22 million), research in this language lags behind languages like Tagalog and Ilokano. Notably, none of the onomatopoeias in any Philippine language have been studied in a manner similar to the present study. Therefore, our findings in Cebuano can provide insights into the contextual and linguistic factors involved in the production and perception of onomatopoeias in Philippine languages and, by extension, other Austronesian languages. Ultimately, our findings aim to contribute to a better understanding of the sound symbolic nature of onomatopoeias more broadly.

Thus, with each research question corresponding to an experimental procedure described above, we ask the following questions:

> **RQ1**: What phonetic changes (vowel quality: F1-F2; pitch: F0; length: duration) do the Cebuano *hagtik* 'click' and *hagtuk* 'knock' onomatopoeias undergo when produced in various imitation contexts?
>
> **RQ2**: How does the perception of the sound source correspond to native and non-native speaker judgments of the distinctive coarticulatory segments of the *click* and *knock* onomatopoeias, i.e., [ik] and [uk] for Cebuano and [ik] and [ok] for Spanish, respectively?

3. Experiment 1 – Production task

In this section, we detailed how we carried out Experiment 1 along with the results. Experiment 1 explored how the Cebuano onomatopoeias *hagtik* 'click' and *hagtuk* 'knock' change phonetically in different imitation contexts. By analyzing vowel quality, pitch, and duration, we found subtle but important acoustic variations depending on context. Generally, the vowels in these onomatopoeias shifted in quality, their pitch tended to increase, and their duration varied, highlighting a clear distinction between neutral and contextualized productions.

3.1. Procedure and Stimuli

The first experiment was a production task participated in by 10 native Cebuano speakers (5 females, 5 males; $M = 21.6$, $sd = 2.1$) from Metro Cebu with normal vision, hearing, and speech faculty. Participants were audio recorded on-site using a Zoom H4n Pro with a Shure SM35 condenser lavalier mic positioned about 2 inches away from the participant's right side of the mouth. The sampling rate for each audio recording was set to 44.1 kHz (16-bit).

Participants were asked to read two sets of randomized stimuli from a computer screen. The first set of stimuli focused on producing *click* contexts (i.e., neutral, mouse click, light switch click, fuse box switch click), and the second set focused on producing *knock* contexts (i.e., neutral, soft door knock, typical door knock, intense door knock). For "neutral" context in particular, we instructed participants to produce an out-of-the-blue sound of *click* and *knock*. Practice sessions were administered before each experimental session to ensure that each participant understood the task.[5] In the end, the stimuli were composed of 4 contexts, 3 vowels /a, i, u/, 3 repetitions, and 2 sets (one for *click* and one for *knock*).

3.2. Acoustic Analysis

Praat (Boersma & Weenink 2023) was used to isolate vowel segments and perform further acoustic analysis. To measure vowel quality, the vowel's first formant (F1) and second formant (F2) frequencies were extracted. F1 indicates vowel height, and F2 indicates backness relative to the vowel

[5] While this instruction can be interpreted in different ways by participants, they seem to have a fairly similar production of what is an out-of-the-blue production of *click* and *knock* in Cebuano based on our results. The inconsistency of their interpretation and, thus, their production, can be seen as a weakness in this approach which may be improved in later iteration of this project. However, by examining multiple acoustic measures, the study will avoid narrowing the effect of context to duration only, for example, i.e., by instructing the participants to lengthen the vowel to indicate changes in the context provided. Examining multiple acoustic measures will hopefully capture the nuances employed by the speakers in their interpretation and production. All of this interpretive and analytical decision aims to capture a better picture of the phonetic changes that occur in the production of onomatopoeias relative to imitative particular contexts.

space. A vowel-intrinsic Bark-transform method was used for normalization to reduce physiological effects. To measure pitch, the fundamental frequency (F0) was extracted. To measure length, the duration of the isolated vowel segment was extracted and recalculated to obtain the vowel duration ratio.

3.3. Statistical Analysis

Statistical analyses for Experiment 1 were completed in R (RStudio Team 2020). The lme4 package was used to evaluate each model. Experiment 1 used dummy-coded linear mixed-effects models with each acoustic measure as the response variable and *context* as the predictor (reference level: "xneut"). Random effects in these models included *speaker* and *repetition*. The emmeans package was used for post-hoc analyses.

3.4. Predictions

Experiment 1 was duly designed to answer RQ1, i.e., "What phonetic changes (vowel quality: F1-F2; pitch: F0; length: duration) do the Cebuano *hagtik* 'click' and *hagtuk* 'knock' onomatopoeias undergo when produced in various imitation contexts?" As such, we predict that Cebuano speakers' production will be sensitive to context in terms of any phonetic changes in their production of *click* and *knock* onomatopoeias. In particular, the vowel quality, pitch, and duration will correspond relatively distinctly to each respective context provided in the experiment (Braver et al. 2016; Samejon 2019).

3.5. Results

3.5.1. Vowel Quality – F1 and F2

Vowels are characteristically known to have identifiable acoustic properties based on the first formant (F1) and the second formant (F2) both of which reflect vocal tract resonance, especially relative to the position of the tongue. F1 values indicate the height (low or high) of the vowel produced, while F2 indicates the backness (or frontedness) of the vowel. Using these two measures, we can track changes in the vowels based on the contexts in which they were produced by the speakers.

For *click* contexts, contextualized /i/ vowel production was lowered and further back compared to neutral production of *click* (Fig. 2A).[6] The model found two significant difference where the fuse context was produced generally with lower /i/ vowel via F1, $\beta = -0.80$, $t = -2.22$, $p < 0.05$, and the mouse context had much backer /i/ vowel via F2, $\beta = 0.98$, $t = 2.79$, $p < 0.01$, relative to neutral contexts. Given the multiplicity of comparisons, we pursued Bonferroni-corrected post-hoc analyses to avoid Type I error. We found that for the F1 and F2 models only the F2 model showed statistically significant comparison between neutral context and mouse context, p<0.05.

For *knock* contexts, the vowel /u/ was also lowered but moved further to the front compared to neutral production of *knock* (Fig. 2B). For F1, we found a couple of significant differences including much lower /u/ for the door knocking and intense door knocking contexts, i.e., $\beta = -0.59$, $t = -2.16$, $p < 0.05$ and $\beta = -0.74$, $t = -2.81$, $p < 0.01$, respectively. However, based on a Bonferroni-corrected post-hoc analysis, the only significant comparison was between neutral context and intense door knocking, $p < 0.05$. For F2, we found a significant difference between the neutral context and a contextualized door knocking, $\beta = -0.82$, $t = -2.54$, $p < 0.05$. Yet this difference is not significant via a Bonferroni-corrected post-hoc analysis, $p = 0.07$.

The results suggested that there were slight but crucial phonetic changes in the vowels of onomatopoeias relative to the contexts in which the speakers produced them. These changes typically manifest via lowering their /i/ and /u/ vowel production, but with a more back /i/ and a more fronted /u/ in most contexts. Significant results point to a more back *click* production and lower *knock* production relative to neutral contexts. In other words, the results suggest a two-way vowel quality distinction between neutral and contextualized production in *click* and *knock* onomatopoeias more generally.

---

[6] Indeed, onomatopoeias cannot be entirely neutral in terms of context given that they are imitations of what speakers experienced in the real world. What *neutral* means in the present study is that the onomatopoeia was 'plainly read' without any additional context given beforehand, compared to others imitation prompts where a particular context was explicitly included.

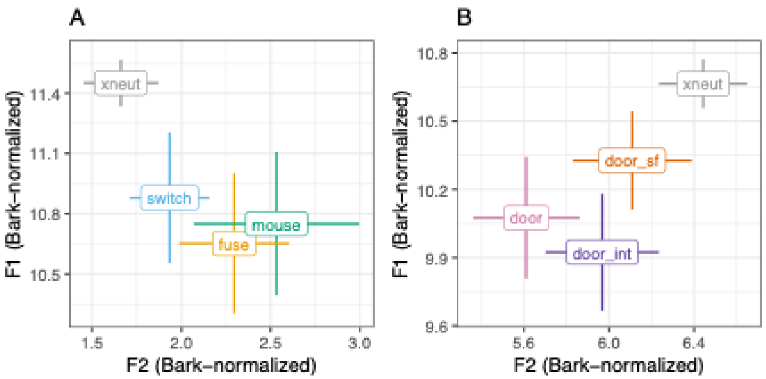

Figure 2. F1 (vertical line) and F2 (horizontal lines) of *click* (A) and *knock* (B) production in various imitation contexts. Lines represent standard error. Each label represents the imitation contexts where "xneut" is the neutral production.

3.5.2. Pitch – F0

Pitch is the highness and lowness of tone in human speech, dependent on the vibrations in the vocal cords. This process is acoustically identifiable based on the speech signal's fundamental frequency (F0) measured in hertz (Hz). Given that pitch is usually anchored on the vowel, we used the vowel's F0 measure. The pitch information from the vowels in our stimuli allowed us to trace whether speakers indicate onomatopoeias in various imitation contexts by changing their pitch.

Across *click* and *knock* contexts, there was a relative change in pitch when the production of onomatopoeias was further contextualized than when no context was specified (Fig. 3). Compared to the neutral context, there was a significant increase in pitch when imitating more contextualized *click* especially for light switches, $\beta = 11.96$, $t = 2.08$, $p < 0.05$. There was also an increase in from a *knock* in neutral context to an intense door knocking, $\beta = 8.62$, $t = 2.65$, $p < 0.01$. However, Bonferroni-corrected post-hoc anlaysis for *click* F0 measure showed no significant comparisons. Similarly, Bonferroni-corrected post-hoc analysis for *knock* F0 measure only showed near significant difference between neutral context and intense door knocking and intense door knock and soft door knock, both at $p = 0.06$.

Similar to vowel quality, changes in pitch in the production of contextualized onomatopoeias were evident. There was a general trend of increased pitch, which distinguishes the neutral context production of onomatopoeias from others. To the extent that there were differences, they were markedly in binary terms.

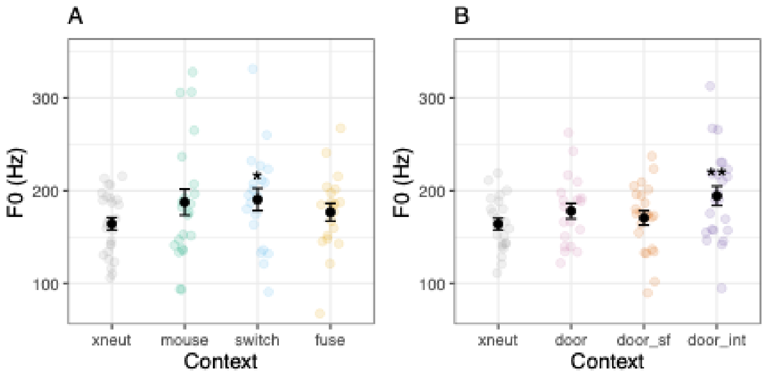

Figure 3. F0 values of *click* (A) and *knock* (B) production in various imitation contexts. Bars represent standard error. Significant difference based on model values: $p < 0.05$ ‘∗’, $p < 0.01$ ‘∗∗’, $p < 0.001$ ‘∗∗∗’.

3.5.3. Length – Duration

The length of a phone refers to the time it takes to produce a particular sound from its onset to offset, and this is usually measured in terms of its phonetic duration (milliseconds). Phonetic duration is used by speakers for emphatic production in Cebuano (Samejon 2019) and other languages like English (Kawahara & Braver 2014). We expanded this observed contextual sensitivity of phonetic duration to the production of *click* and *knock* onomatopoeias. In particular, we considered the phonetic duration ratio of the vowels /i/ and /u/ in various imitation contexts. The decision to use duration ratio instead of raw duration values was made to lessen the effect of speech rate on vowel duration. At the same time, duration ratio better shows how much of the vowel's duration has changed in various imitation contexts relative to the total duration of the word.

There was increased duration ratio in *click* contexts, where the duration of /i/ vowel increased relative to the duration of the word. However, speakers only distinguished this between their neutral production of *click* and their production of a mouse click and nothing else (Fig. 4A). There was at least 5 percent increase in the duration of /i/ vowel in a word in contrast to their typical production of clicking sound, $\beta = 5.04$, $t = 4.61$, $p < 0.01$. Bonferroni-corrected post-hoc anlaysis, on the other hand, showed three significant comparisons, i.e., between neutral context and mouse click, mouse click, and fuse click, and mouse click and light switch, all at $p < 0.01$. This three-way difference is surprising, i.e., neutral context > mouse click > fuse click/light switch click.

In *knock* contexts, the duration ratio of /u/ vowel was significantly increased in all contexts (Fig. 4B). Relative to their neutral production of *knock*, the /u/ vowel in the context of door knocking increased to about 4 percent, soft door knocking increased to 6.11 percent, and intense door knocking to 5.67 percent, i.e, $\beta = 4$, $t = 2.56$, $p < 0.05$, $\beta = 6.11$, $t = 4.05$, $p < 0.01$, and $\beta = 5.67$, $t = 3.72$, $p < 0.01$, respectively. When examining these multiple comparisons further via Bonferroni-corrected post-hoc analysis, the significant difference across *knock* contexts were between neutral context and soft door knocking, $p < 0.01$, and between neutral context and intense door knocking, $p < 0.01$.

The duration of the key distinguishing vowels in both *click* and *knock* onomatopoeias showed contextual effect. For *click* contexts, we only found one context where /i/ vowel duration ratio increased while all *knock* contexts showed increase in /u/ vowel duration relative to the neutral production. Overall, we found a general two-way distinction between neutral and contextualized production, and across contextualized productions of onomatopoeias relative to duration.

4. Experiment 2 – Production Task

In this section, we described how we prepared and administered Experiment 2, and reported the results of the experiment. Experiment 2 examined how native Cebuano speakers perceive the association between sound sources and onomatopoeias using a listening-perception task with rating scales. What we found is that the Cebuano participants generally rated the expected *click* and *knock* sounds in Cebuano stimuli similarly to a previous study. Notably, *click* was associated with both [uk] and [ik], while *knock* was strongly linked to [uk]. Yet, when listening to Spanish stimuli, the Cebuano participants did not replicate these results.

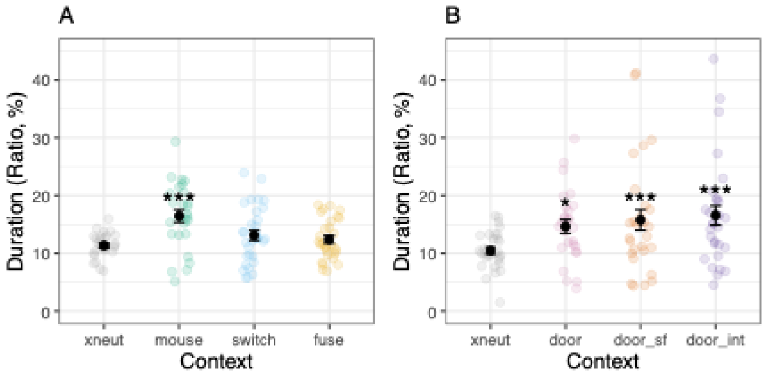

Figure 4. Duration ratio of *click* (A) and *knock* (B) production in various imitation contexts. Bars represent standard error. Significant difference based on model values: $p < 0.05$ '*,' $p < 0.01$ '**,' $p < 0.001$ '***.'

4.1. Procedure and Stimuli

The second experiment was a perception task participated in by 28 native Cebuano speakers ($M = 29.1$, $sd = 6.4$), none of whom were participants in Experiment 1. Instead of visual stimuli, participants were presented with several audio stimuli online via Qualtrics. Each participant was asked to confirm that they knew what a typical mouse click sounds like when being tested on *click* associations and what a typical door knock sounds like when being tested on *knock* associations. Based on the stimuli they heard, they had to rate how closely the sound they heard was associated with their knowledge and experience of a typical mouse click or door knock, with "1" indicating no association and "10" indicating a perfect association.

The stimuli used in the study were clipped audio files from *hagtik*, *hagtak*, and *hagtuk*, i.e., coarticulated [ik], [ak], and [uk] segments, produced by four native Cebuano speakers (2 females) from Experiment 1. To process their audio recordings, the vowel was first located, and the vowel duration was divided into four parts in Praat (Boersma & Weenink 2023). Then, to isolate the coarticulated [ik], [ak], and [uk] segments, the onset of the final quarter of the vowel became the starting point, extending the total duration of the segments rightward to approximately 100 milliseconds (Fig. 5). We expressed this clipping procedure mathematically as:

$$y = \left(\frac{3x}{4}\right) + m$$

where $3x/4$ is the starting point of the final quarter of the partitioned vowel sound, m is the time added to the result of the partitioned vowel duration, and y is the duration of the experimental token from onset to offset which should be around 100 milliseconds in total.

Next, these isolated coarticulated segments were extracted and again processed in Praat to normalize their intensity to 73 dB, which controlled for loudness of each clipped audio files. Lastly, these clipped and processed audio files were imported to Qualtrics as perception experiment stimuli. The Cebuano audio stimuli created and used for this project were hosted in https://osf.io/n9u5e/.

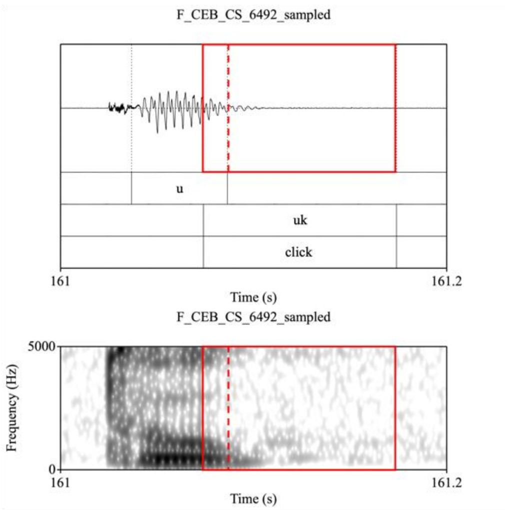

Figure 5. Example of the clipping procedure for the perception task stimuli. The smaller jagged red rectangle marks the final quarter of the isolated vowel. The larger solid red rectangle marks the total duration of the stimuli. The first tier labels the isolated vowel. The second tier labels the coarticulated segment. The third tier labels the sound association the participants are to make with the audio clip.

In order to test native and non-native perception, we included the stimuli in Cebuano and the publicly available coarticulation stimuli in Spanish by Assaneo et al. (2011). The audio clips from Assaneo et al. (2011) were also processed to normalize their intensity to 73 dB before being used in the present study. A total of 4 sets of stimuli were created, including 32 audio clips, i.e., 2 sets for each language, with 8 audio clips per set. Two groups of participants listened to the same 4 sets of stimuli: one group rated how much the sounds they heard resembled a typical *click*, and the other group did the same for a typical *knock* based on their experience and knowledge of these sounds.

Finally, an important step in the perception task was that participants were not informed that the sounds they would hear came from speech signals or that they were drawn from two different languages, namely Cebuano and Spanish. Participants were simply instructed to rate whether the 'sound' they heard resembled the sound of clicking or knocking. This approach aimed to allow participants to focus on associating their knowledge and experience of *click* and *knock* sounds to the stimuli they heard, rather than feeling pressured to provide a correct response.

4.2. Statistical Analysis

Similar to Experiment 1, statistical analyses for Experiment 2 were also completed in R (RStudio Team, 2020). Given the rating-type response indicated by participants, the ordinal package was used to evaluate each model. Experiment 2 used cumulative link mixed-effects models with *rating* for each language's *click* and *knock* perception as the response variable and coarticulated segment as the predictor (reference level: [ak]). Random effects in these models included *speaker*.[7]

4.3. Predictions

Experiment 2 was pursued to provide an answer for RQ2, i.e., "How does the perception of the sound source correspond to native and non-native speaker judgments of the distinctive coarticulatory segments of the *click* and *knock* onomatopoeias, i.e., [ik] and [uk] for Cebuano and [ik] and [ok]

[7] We would like to thank the anonymous reviewer who recommended to use cumulative link mixed-effects models instead of mixed-effects linear regression models, given the scale participants used for rating the audio stimuli, i.e., 1-10.

for Spanish, respectively?" We predict that Cebuano speakers will be able to associate the most identifying coarticulated segments of Cebuano *click* and *knock* onomatopoeias, i.e., [ik] and [uk], more favorably with their knowledge and experience of the sound of clicking or knocking in the real world. Moreover, the crosslinguistically identifiable coarticulated segments of *click* and *knock* are acoustically optimized in imitative speech (Assaneo et al. 2011). This suggests that a speaker will be able to easily identify a click sound with the coarticulated [ik] segment from another language. Therefore, we also predict that the judgment of Cebuano speakers on similar *click* and *knock* coarticulated segments from another language will pattern in the same way as the pattern found in Spanish by Assaneo et al. (2011).

4.4. Results

4.4.1. Cebuano Stimuli Perception

The native Cebuano speaking participants demonstrated clear associations between *click* and *knock* relative to the coarticulated segments when they listened to the Cebuano stimuli (Fig. 6A).

For *click*, [ak] was not strongly associated to the sound of *click* based listener ratings ($M = 3.14$, $sd = 2.20$) while [ik] and [uk] were both rated much higher ($M = 4.20$, $sd = 2.88$; $M = 4$, $sd = 2.45$). The increased rating of [ik] and [uk] relative to [ak] were also statistically significant, $\beta = 0.52$, $z = 2.78$, $p < 0.01$, and $\beta = 0.41$, $z = 2.21$, $p < 0.05$. To investigate which responses were robustly different given the multiplicity of comparisons, a Bonferroni-corrected post-hoc analysis was pursued and returned only a significant difference between [ak] and [ik], $p < 0.05$. We, nonetheless, found a near significant difference between [ak] and [uk], $p = 0.08$.

For *knock*, [uk] was rated much higher as the sound associated to *knock* ($M = 5.40$, $sd = 3.10$) compared to [ak] and [ik] ($M = 4.18$, $sd = 2.72$; $M = 3.50$, $sd = 2.50$). The increased rating for [uk] relative to [ak] was statistically significant, $\beta = 1.23$, $z = 3.11$, $p < 0.01$. A Bonferroni-corrected post-hoc analysis for multiple comparisons was also pursued and revealed that both [ak] and [ik] were significantly different from [uk], both at $p < 0.01$.

In general, the perception results revealed that native Cebuano speakers rated the expected coarticulated segments in the Cebuano stimuli, patterning similarly to Assaneo et al. (2011)'s results. The results also showed an emerging association of *click* to [uk] in addition to [ik], while *knock* was only strongly associated with [uk]. This expanded association of [uk] with *click* might not be solely due to acoustic similarity but rather to a more conventionalized use of similar onomatopoeias in Cebuano. We explore this further in the discussion section.

4.4.2. Spanish Stimuli Perception

Results for the Spanish stimuli perception revealed that the native Cebuano participants had difficulty in associating non-native sounds to either *click* and *knock* (Fig. 6B).

In fact, [ek] and [ik] were the lowest rated sound to be associated to *click*, i.e., $M = 1.86$, $sd = 1.21$ and $M = 2.04$, $sd = 1.14$, respectively. We also found that [ok] is relatively the highest rated association to *click*, $M = 2.64$, $sd = 1.21$. The difference between the rating given to [ak] and the expected high front vowels, however, was not significant, [ek]: $\beta = -0.48$, $z = -1.68$, $p = 0.09$, [ik]: $\beta = -0.17$, $z = -0.64$, $p = 0.53$.

Across the coarticulated segments in Spanish, [ok] was highest rated sound to be associated to the *knock*, $M = 4.29$, $sd = 3.12$. The increased rating for [ok] relative to [ak], however, was not significant, $\beta = 0.39$, $z = 1.58$, $p = 0.11$.

The native Cebuano participants failed to replicate the results in Assaneo et al. (2011) when listening to the production of *click* and *knock* coarticulated segments in the Spanish stimuli. These results were surprising given that *click* and *knock* sound sources are very similar across languages by virtue of their onomatopoeic representations. Yet the supposed stable speech sounds in *click*, i.e., [ik], and *knock*, i.e., [ok], onomatopoeias were not perceptually different from other alternatives for the Cebuano listeners.

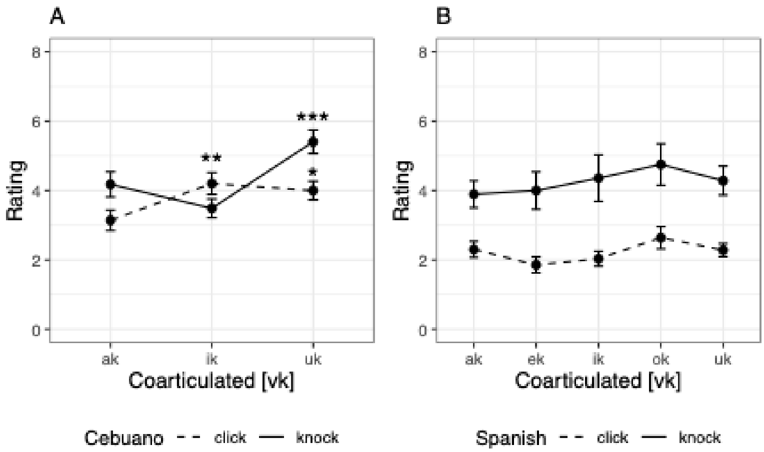

Figure 6. Rate of association by native Cebuano speakers on *click* and *knock* coarticulated segments when listening to Cebuano stimuli (A) and Spanish stimuli (B). Compare this Figure 1. Bars represent standard error. Significant difference based on model values: $p < 0.05$ '∗', $p < 0.01$ '∗∗', $p < 0.001$ '∗∗∗'.

5. General Discussion

The present findings revealed that the production of *click* and *knock* onomatopoeias is context-sensitive and follows a language-dependent perceptual association to the sounds they imitate. These findings improve our understanding of the phonetic changes involved in the contextualized production of onomatopoeias and the language constraints that determine sound-to-meaning associations of coarticulated segments. Moreover, the predictions set forth in the study were not entirely borne out. Inasmuch as there is context-sensitive production of onomatopoeias, it has mainly been a two-way distinction and can be attributed to the common linguistic tendencies in the Cebuano language. Interestingly, while crosslinguistic evidence supports a strong association between the onomatopoeias and the sounds they imitate, perception of this association may be language-specific.

5.1. Two-Way Distinction

A surprising finding is that a two-way distinction persisted in vowel quality. This may be a type of change where, in order to create contrasts, vowels are lowered and centered in contextual imitative production relative to a more neutral one. Proposed explanations for changes in vowel quality in Cebuano have been made, i.e., either stress-unstressed distinction (Samejon 2022; see also Garellek & White 2015) or syllable/phrase-finality (Tanangkingsing 2009; see also Kaufman forthcoming). Neither explanation was satisfactory based on the results. Since the Cebuano *click* and knock *vowels* investigated were found on the final syllable of the onomatopoeia, which is lexically unstressed, there would have been no change in vowel quality in their production. In other words, these vowels will have remained in their 'neutral' state because there is no stress-shifting morpheme that was added to the word (Shryock 1993)–but this had not been the case. The quality of the vowel still changed. On the other hand, syllable-finality cannot be a better explanation because it predicts that the /i/ and /u/ are lowered to /e/ and /o/ in the final position. Since the vowels in the onomatopoeias were further lowered and then centralized, it might be improbable that the already lowered vowels à la Tanangkingsing are produced as something like a schwa or [a]. An alternative explanation to what we have observed may perhaps be the effect of *context*. When linguistic expectations fail, speakers rely on the richness of context to establish the acoustic distinctions necessary to communicate a particular utterance (Kawahara & Braver 2014), or in the case of onomatopoeias, to imitate extralinguistic sounds in context.

Similar to vowel quality, pitch differences suggest a two-way distinction. Cebuano is known to have only low-high intonation or pitch differences, usually found at phrase edges (Tanangkingsing 2009; Wolff 1972). While the neutral production was not at the phrase edges, the production of contextualized *click* and *knock* is in citation form, which might have exacerbated the difference between neutral production and production in other contexts. This is a potential weakness in the production stimuli used in the study. Additionally, upon further reflection on pitch, our results suggest that there may be no significant difference regardless of context. This implies that pitch may be less important for distinguishing various contexts of onomatopoeic productions. Analyzing pitch in Cebuano requires more attention and can be improved on future work.

For length, Ladefoged & Maddieson (1996) observed this crosslinguistic tendency in lexical contrasts, at least for durational differences. The two-way distinction for duration in this study is expected because Samejon (2019) observed this phonetic behavior in Cebuano much earlier. This does not necessarily mean that languages with two-way distinctions cannot imagine, exert effort, or even produce more gradient distinctions (Braver et al., 2016), but a two-way distinction is a default or the strongest tendency for these languages more generally.

It is reasonable then to ask why a two-way distinction is preferred. Perhaps, a two-way distinction may be less perceptually difficult compared to a three-way distinction, for example. Liljencrants & Lindblom (1972) theorized that in order to minimize confusability, vowels disperse in locations of the vowel space far enough from each other to maximize vowel distinctiveness. The same mechanism may be at work in the contextualized production of onomatopoeias, i.e., the pervasiveness of a perceptually easier but still contrastive two-way distinction across the board. Moreover, pre-existing language-specific distinctions in vowel quality, pitch, and length may have contributed to speakers' predisposition for a two-way distinction rather than a more gradient distinction. Such preponderance of a two-way distinction in phonological features of languages like Cebuano may warrant closer examination in light of this binary distinction classically espoused in the literature (e.g., whether a sound is +voiced or -voiceless, Chomsky & Halle 1968; but see three-way length distinction in Mixe, Hoogshagen 1959).

5.2. Crosslinguistic *Click* and *Knock*

The linguistic structure of *click* and *knock* onomatopoeias in the languages of the world are grossly similar, yet the perceptual association of each one to their respective sound sources differ. Across languages, imitation of *click* and *knock* sounds is composed of at least two identifying coarticulated segment [ik] for *click* and [ok] for *knock*. Assaneo et al. (2011) established that [ik] and [ok] were the closest linguistic representation of the acoustic characteristics of real world *click* and *knock* sounds, which suggests a strong sound symbolic and transparent association between sound and meaning. The native Cebuano speaking group in the study, however, did not follow this generalization especially when listening to Spanish coarticulated segments. Instead, the native Cebuano speaking group in the study only exhibited similar patterning as Assaneo et al. (2011)'s when listening to Cebuano coarticulated segments.

Also, the current findings have implications for the phoneme inventory similarity hypothesis, which posits that linguistic patterns are similar between languages with the same phoneme inventory (Lavoie 2002; Samejon 2021). Cebuano does not have the phoneme /x/, so the decision to replace the velar plosive /k/ with velar fricative /x/ in the final consonant of the coarticulated segments in Assaneo et al. (2011) may have introduced difficulty in the perception of the Spanish stimuli by Cebuano speakers. This alternation between /k/ and /x/ should not have been a problem for Cebuano speakers because the acoustic properties of word-final /k/ and /x/ in the onomatopoeias in question were found to be analogous to each other (Assaneo et al. 2011). Yet, when presented with a non-native final consonant in /x/, the sound symbolic valence of the Spanish stimuli deteriorated fairly quickly and contributed to Cebuano speakers' failure to associate the expected coarticulated segments to *click* and *knock* sounds.

The use of either a velar plosive or fricative as the final sound may have an effect on the noise-to-duration ratio of the stimuli. One may even argue that this exacerbated the difference in the perception of Cebuano speakers on the Spanish set of the stimuli. Indeed, while the present study's stimuli were about 100 ms each, each stimulus mostly contained silence due to the nature of unreleased /k/ in the final coarticulatory segment of Cebuano *click* and *knock* onomatopoeias. This means that only a quarter of the stimuli, or 25 ms, contained acoustic information and not silent compared to the 100 ms of non-silent audio in the Spanish stimuli. Had Cebuano speakers found it problematic to associate what they heard from a mostly silent stimulus, they would have failed in the Cebuano onomatopoeia perception task instead of the Spanish one. This means that perceptual success among Cebuano speakers do not entirely rely on whether they hear a sound within the full 100 ms or not. Their success seems to rely mainly on the meaningfulness or familiarity of the sound they hear even when if it is only about 25 ms, i.e., successfully understanding the information contained in the stimuli that used the velar plosive compared to the stimuli with velar fricative regardless of the length of acoustic signal.

More crosslinguistic work is needed, but it is more likely that the non-distinction among Cebuano speakers is due to crosslinguistic differences rather than the duration of acoustic information contained in the sampled Spanish stimuli. Besides, relative to the duration, humans are sensitive to durational cues as short as 5 ms to 12.5 ms (Näätänen et al. 1989; Johnson 2010). Adding to the fact that the coarticulated segments of *click* and *knock* onomatopoeias are perceptible, the difference in the noise-to-duration ratio in the Cebuano stimuli becomes irrelevant to the perceptual asymmetry found in the performance of Cebuano speakers on Cebuano and Spanish *click* and *knock* sound associations.

Lastly, the perceptual difference between *click* and *knock* onomatopoeias can go beyond their typical crosslinguistic phonetic tendencies. In Cebuano, we observe a very typical [ik] 'click' and [uk] 'knock' distinction, but Cebuano speakers' perception of *click* was also high for coarticulated [uk] segments. This was surprising given that Cebuano onomatopoeias *hagtik* and *hagtuk* are contrastive relative to *click* and *knock* onomatopoeias. However, looking further into the possible alternatives in the language, there is another word that indicates a clicking event: *tupluk-tupluk*, which is a word typically associated to typing on a keyboard. Using a digital or physical keyboard is a more recent activity and speakers have to deal with the fact that at some point they will have to communicate the sound that clicking, tapping, or pressing on a keyboard produces. Among Cebuano speakers, they seem to agree to having the final coarticulated [uk] segment represent the clicking sound. In fact, Tagalog and Czech have similar sounds for typing on a keyboard, too, i.e., *ku chuk* and *t'uk t'uk*, respectively, which makes the Cebuano case not too isolated. The connection between the representation and perception of Cebuano and Czech *click* onomatopoeia, two different languages families, may be explored further as well. Moreover, the word *tupluk-tupluk* in Cebuano is not necessarily sound imitative but may have been conventionalized to extend and carry the same imitative function as the Cebuano *click* onomatopoeia. Conventionalization of a word that describes an event into an onomatopoeia is not an uncommon process in onomatopoeic word formation crosslinguistically, and that eventive word may have its roots on imitation in the first place anyway (Körtvélyessy & Štekauer 2024; see also Perlman et al. 2015). As such, Cebuano speakers became comfortable in associating both [ik] and [uk] when perceiving clicking sounds.

6. Conclusion

The present study demonstrated that the production and perception of onomatopoeias are reflexes of crosslinguistic and language-specific constraints. Crosslinguistically, Cebuano exhibited the widely attested two-way or binary distinction when establishing contrasts on contextualized *click* and *knock* onomatopoeias. Language-specifically, Cebuano speakers found it challenging to perceive non-native *click* and *knock* onomatopoeias and even extended the use of coarticulated [uk] segments for clicking sounds (in addition to [ik]) when perceiving Cebuano clicking sound. Perhaps, the collective experience of speakers to keyboard typing made the sound symbolic nature of coarticulated [uk] sound conventionalized for clicking as well.

Additionally, the typological features and sociocultural setting of the Cebuano language and its speakers resemble some Austronesian and Philippine ethnolinguistic groups. Thus, the findings of the present study may potentially provide a general prediction on the patterns of sound-meaning associations for these groups of language speakers. Chomsky and colleagues (Hauser et al. 2002) have also already noted that humans' sound imitation through speech is somewhat overlooked. Consequently, our findings will further contribute to a more systematic understanding of sound symbolic imitations in the world's languages.

References

Assaneo, María Florencia, et al. “The Anatomy of Onomatopoeia.” PLoS ONE, vol. 6, no. 12, 2011, e28317, doi:10.1371/journal.pone.0028317.

Berent, Iris. “Are Phonological Representations of Printed and Spoken Language Isomorphic? Evidence from the Restrictions on Unattested Onsets.” *Journal of Experimental Psychology*, vol. 34, no. 5, 2008, pp. 1288–304, doi:10.1037/0096-1523.34.5.1288.

Boersma, Paul, and David Weenink. *Praat: Doing Phonetics by Computer [Computer Program]*. 2023, www.praat.org/.

Braver, Aaron, et al. “The Phonetics of Emphatic Vowel Lengthening in English.” *Proceedings of the Annual Meetings on Phonology*, 2016, pp. 1–12, doi:10.3765/amp.v2i0.3754.

Bril, Isabelle. “Onomatopoeia in Northern Amis.” *Onomatopoeia in the World’s Languages: A Comparative Handbook*, edited by Lívia Körtvélyessy and Pavol Štekauer, De Gruyter Mouton, 2024, pp. 805–16, doi:10.1515/9783111053226-067.

Childs, George T. “Sound Symbolism.” *The Oxford Handbook of the Word*, edited by John R. Taylor, Oxford University Press, 2015, pp. 284–302, doi:10.1093/oxfordhb/9780199641604.013.030.

Chomsky, Noam, and Morris Halle. *The Sound Pattern of English*. Harper; Row, 1968, web.mit.edu/morrishalle/pubworks/papers/1968_Chomsky_Halle_The_Sound_Pattern_of_{E}nglish.pdf.

Garellek, Marc, and James White. “Phonetics of Tongan Stress.” *Journal of the International Phonetic Association*, vol. 45, no. 1, 2015, pp. 13–34, doi:10.1017/S0025100314000206.

Güldemann, Tom. *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*. Mouton de Gruyter, 2008, doi:10.1515/9783110211450.

Hauser, Marc D., et al. “The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?” *Science*, vol. 298, no. 5598, 2002, pp. 1569–79, doi:10.1126/science.298.5598.1569.

Haynie, Hannah, et al. “Sound Symbolism in the Languages of Australia.” *PLoS One*, vol. 9, no. 4, 2014, p. e92852, doi:10.1371/journal.pone.0092852.

Hockett, Charles F. “In Search of Jove’s Brow.” *American Speech*, vol. 53, no. 4, 1978, pp. 243–313, www.jstor.org/stable/455140.

Hoogshagen, Searle. “Three Contrastive Vowel Lengths in Mixe.” *STUF-Language Typology and Universals*, vol. 12, no. 1-4, 1959, pp.

111–15, doi:10.1524/stuf.1959.12.14.111.

Intlekofer, Darlene, and Jason Bishop. "The Role of Prosody in Conditioning Tagalog o/u Variation." *Proceedings of Speech Prosody 2016*, 2016, pp. 1105–08.

Jespersen, Otto. "Symbolic Value of the Vowel /i/." *Linguistica. Selected Papers in English, French and German by Otto Jespersen*, 1933, pp. 283–303, www.taylorfrancis.com/chapters/mono/10.4324/9780203857199-17/symbolic-value-vowel-1933-otto-jespersen.

Johansson, Niklas Erben, et al. "The Typology of Sound Symbolism: Defining Macro-Concepts via Their Semantic and Phonetic Features." *Linguistic Typology*, vol. 24, no. 2, 2020, pp. 253–310, doi:10.1515/lingty-2020-2034.

Johnson, Keith. *Acoustic and Auditory Phonetics*. 3rd ed., Wiley-Blackwell, 2010.

Kaufman, Daniel. "The Languages of Central and Southern Philippines." *Oxford Guide to the Malayo-Polynesian Languages of Southeast Asia*, edited by Antoinette Schapper and Sander Adelaar, Oxford UP, forthcoming.

Kawahara, Shigeto, and Aaron Braver. "Durational Properties of Emphatically Lengthened Consonants in Japanese." *Journal of the International Phonetic Association*, vol. 44, no. 3, 2014, pp. 237–60, doi:10.1017/S0025100314000085.

Köhler, Wolfgang. *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*. W. W. Norton & Company, 1970.

Körtvélyessy, Lívia. "Onomatopoeia–a Unique Species?" *Studia Linguistica*, vol. 74, no. 2, 2020, pp. 506–51, doi:10.1111/stul.12133.

Körtvélyessy, Lívia, and Pavol Štekauer. "Introduction: Why Onomatopoeia?" *Onomatopoeia in the World's Languages: A Comparative Handbook*, edited by Lívia Körtvélyessy and Pavol Štekauer, vol. 10, De Gruyter Mouton, 2024, pp. 1–22, doi:10.1515/9783111053226.

Kwon, Nahyun, and Erich R. Round. "Phonaesthemes in Morphological Theory." *Morphology*, vol. 25, 2015, pp. 1–27, doi:10.1007/s11525-014-9250-z.

Ladefoged, Peter, and Ian Maddieson. *The Sounds of the World's Languages*. 2nd ed., Blackwell, 1996.

Lavoie, Lisa M. "Subphonemic and Suballophonic Consonant Variation: The Role of the Phoneme Inventory." *ZAS Papers in Linguistics*, vol. 28, 2002, pp. 39–54, doi:10.21248/zaspil.28.2002.158.

Liljencrants, Johan, and Björn Lindblom. "Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast." *Language*, vol. 48, 1972, pp. 839–62, doi:10.2307/411991.

Näätänen, Risto, et al. "Do Event-Related Potentials Reveal the Mechanism of the Auditory Sensory Memory in the Human Brain?" *Neuroscience Letters*, vol. 98, no. 2, 1989, pp. 217–21, doi:10.1016/0304-3940(89)90513-2.

Nuckolls, Janis B. "The Case for Sound Symbolism." *Annual Review of Anthropology*, vol. 28, no. 1, 1999, pp. 225–52, www.jstor.org/stable/223394.

Ohala, John J. "The Frequency Code Underlies the Sound-Symbolic Use of Voice Pitch." *Sound Symbolism*, edited by Leanne Hinton et al., Cambridge UP, 1994, pp. 325–47, www.cambridge.org/core/books/abs/sound-symbolism/frequency-code-underlies-the-soundsymbolic-use-of-voice-pitch/1D05AB5A9266731D6C66FA7BD6F4058C.

Perlman, Marcus, et al. "Iconicity Can Ground the Creation of Vocal Symbols." *Royal Society Open Science*, vol. 2, no. 8, 2015, p. 150152, doi:10.1098/rsos.150152.

Ramachandran, Vilayanur S., and Edward M. Hubbard. "Synaesthesia–a Window into Perception, Thought and Language." *Journal of Consciousness Studies*, vol. 8, no. 12, 2001, pp. 3–34, www.sfu.ca/~kathleea/colour/docs/Ram&Hub_2001.pdf.

RStudio Team. *RStudio: Integrated Development for R*. RStudio, PBC, 2020, www.rstudio.com/.

Samejon, Kevin. "Emphatic Vowel Length in Cebuano." *Journal of Philippine Linguistics*, vol. 50, 2019, pp. 18–32, www.pjl-phil.com/article/2019/50/kevin-samejon.

———. *Stress-Conditioned Vowel Change in Reduplicated and Suffixed Words: Evidence from Cebuano /u/*. 2022, static.sched.com/hosted_files/seals31/a0/Samejon_Cebuano_SEALS2022.pdf.

———. *Temporal Nasalization Patterns Across Word Boundaries*. QP. Boston University, 2021.

Sapir, Edward. "A Study in Phonetic Symbolism." *Journal of Experimental Psychology*, vol. 12, no. 3, 1929, p. 225, doi:10.1037/h0070931.

Saussure, Ferdinand. *Course in General Linguistics*. Edited by Albert Sechehaye and Charles Bally, Translated by Wade Baskin, The Philosophical Library, 1959.

Shinohara, Kazuko, and Shigeto Kawahara. "A Cross-Linguistic Study of

Sound Symbolism: The Images of Size." *Annual Meeting of the Berkeley Linguistics Society*, 2010, pp. 396–410, doi:10.3765/bls.v36i1.3926.

Shryock, Aaron. "A Metrical Analysis of Stress in Cebuano." *Lingua*, vol. 91, no. 2–3, 1993, pp. 103–48, doi:10.1016/0024-3841(93)90010-T.

Tanangkingsing, Michael. *A Functional Reference Grammar on Cebuano*. National Taiwan University, 2009.

Winter, Bodo, et al. "Which Words Are Most Iconic? Iconicity in English Sensory Words." *Interaction Studies*, vol. 18, no. 3, 2017, pp. 443–64, doi:10.1075/is.18.3.07win.

Wolff, John Ulrich. *A Dictionary of Cebuano-Visayan*. Cornell University, 1972, www.gutenberg.org/files/40074/40074-h/40074-h.htm.

Zuraw, Kie. "Using the Web as a Phonological Corpus: A Case Study from Tagalog." *Proceedings of the 2nd International Workshop on Web as Corpus (EACL06)*, edited by Adam Kilgarriff and Mark Baron, 2006, pp. 59–66.

Address for Correspondence

Kevin Samejon
Department of Linguistics
Boston University
621 Commonwealth Ave.
Boston, MA 02215
United States of America

ksamejon@bu.edu

Submitted Date: July 15, 2024
Accepted Date: December 8, 2024