Japanese speakers can infer specific sub-lexicons using phonotactic cues*

Shigeto Kawahara[1] and Gakuji Kumagai[2]
[1]Keio University/International Christian University and
[2]Kansai University

Abstract

Phonotactic restrictions do not usually hold uniformly across the entire lexicon of a particular language, and thus the lexicon of a natural language is often assumed to consist of different sub-lexicons. A question that arises is how specific these sub-lexicons can be. A classic, conservative approach is to posit only broad distinctions, such as the distinction between native words vs. borrowed words. An alternative approach is to posit more specific miniature lexicons, such as a set of morphemes that show a particular morpheme-specific pattern or a set of loanwords from a particular language. With this general theoretical issue in mind, this paper first points out that there are phonotactic restrictions that can cue a very specific class of words in Japanese; e.g. geminate /ɾɾ/ occurs almost exclusively in loanwords from Italian. Building upon these novel observations, the current experiments tested whether Japanese speakers can infer specific word classes such as "snack names", "Italian restaurant names" and "German names" based on particular phonotactic cues. The results of the two experiments support the idea that the lexicon of a natural language can consist of very specific sub-lexicons, at least going beyond the often posited native vs. loanword distinction, and that these specific sub-lexicons can be psychologically real.

Keywords: sub-lexicon, Japanese, the Pocky effect, geminate /ɾɾ/, geminate /hh/

*We would like to thank Hiro Katsuda for his advice on using the BCCWJ to confirm the initial intuitions we had, and Tomoko Tanaka for her help with the analyses of the corpus. We also received helpful feedback on the design of the experiments from the participants of a seminar that the first author gave at International Christian University, especially Honoka Endoh, Rina Furusawa and Haseru Ida. We received constructive comments from two anonymous reviewers, Jason Shaw and the audience at the Phonological Society of Japan, which improved the quality of the paper. All remaining errors are ours.

1 Introduction

1.1 General theoretical background

Phonotactic restrictions—restrictions on which sounds can appear in what position and how these sounds can be arranged—do not often hold uniformly across the entire lexicon of a given language. Typical cases come from differences between native words and loanwords, in which particular sounds and/or sound sequences are often only allowed in loanwords but not in native words. For instance, in English most if not all words with stress on their final syllables are loanwords from French (Gelbart 2005). Japanese also exhibits such patterns, in which the lexicon is considered to consist of several lexical strata: native words, Sino-Japanese words, recent loanwords and mimetic words (Ito & Mester 1995, 1999, 2008). These sub-lexicons are subject to different sets of phonotactic restrictions; e.g. voiced obstruent geminates are allowed only in recent loanwords but not in other types of words; similarly, Sino-Japanese morphemes and mimetic roots are maximally bimoraic (Ito & Mester 1996), but no such size restrictions appear to hold in native words or recent loanwords. See Ito & Mester (1995) for a review of similar examples from other languages.

Experimental studies have shown that such phonotactic restrictions that hold specifically to a subset of the lexicon are, at least in some cases, psychologically real, in that they can influence the perception of ambiguous acoustic signals—listeners can use these cues to decipher the lexical class of the stimuli (Gelbart 2005; Gelbart & Kawahara 2007; Moreton & Amano 1999). For instance, Sino-Japanese does not allow either singleton /p/ or long /aa/, whereas loanwords allow both. When Japanese listeners hear a nonce word with /p/, they tend to perceive the /a/-/aa/ continuum more likely to be long than the control condition (Moreton & Amano 1999). On the other hand, recent loanwords generally do not exhibit a palatalized /ɾj/, which is very common in Sino-Japanese compounds (e.g. /rjokoo/ 'travel'). Upon hearing nonce words with a palatalized /ɾj/, Japanese listeners are biased toward perceiving the /a/-/aa/ continuum to be short compared to the control condition. This experiment by Moreton & Amano (1999) thus shows that listeners can use phonotactic cues—such as singleton /p/ and palatalized /ɾj/—to decipher which lexical class the incoming acoustic signal belongs to, which in turn affects their short /a/ vs. long /aa/ perception.

One question that arises in this context is how fine-grained the sub-division of the lexicon can be in natural languages. The "classic" (and also conservative) view is to postulate only very general divisions, perhaps with independent etymological motivations. Take the case of Japanese, for example—Ito & Mester (1995, 1999, 2008) posit four general strata (native, Sino-Japanese, foreign and mimetics). Similarly, in English, we could posit a very general distinction between native words and recent loanwords (Gelbart 2005), and perhaps within "native" words, a distinction between Germanic roots vs. Latinate roots (Chomsky & Halle 1968).

On the other hand, there is an alternative, less-conservative proposal. For example, the research approach now widely known as "co-phonology" argues that there can co-exist many types of morphologically-conditioned phonological patterns in a single language, and posits that there can be as many number of phonological sub-systems as the number of such morpheme-specific patterns (Anttila 2002; Inkelas et al. 1996, 1997; Inkelas & Zoll 2007; Orgun 1996; Sande 2020 among many others—see also Pater 2005). A similar idea is proposed by a line of work dubbed "sublexical phonology" (Becker & Gouskova 2016; Blake & Becker 2015; Gouskova & Ahn 2024; Gouskova et al. 2015), which posits that "learning lexically specific morphological and phonological rules involves separating the lexicon into sublexicons. Phonological generalizations about the application of such rules are encoded in part as phonotactic grammars learned over sublexicons." (Gouskova & Ahn 2024: 6).[1]

With this general theoretical debate in mind, the current study explores how specific such sublexicons can be. Our case study is based on novel observations about the phonotactic restrictions that are associated with very specific parts of the Japanese lexicon. The experiments reported below examined whether Japanese speakers are only sensitive to the broad sub-lexicons that have been traditionally posited (i.e. native, Sino-Japanese, foreign and mimetics), or alternatively, whether they are sensitive to more fine-grained distinctions. This question is addressed through the investigation of the three novel, specific phonological observations, explained in detail in the next subsection.

Before we proceed to the specifics, one general remark is in order. Japanese offers a particularly interesting testing ground to address this general issue, because not only are the traditional distinctions between native words, Sino-Japanese words and recent loanwords clearly motivated on phonological grounds (Ito & Mester 1995, 1999, 2008), Japanese uses different orthographic systems for different word classes; i.e. the *hiragana* system for native words, the *katakana* system for loanwords and mimetics, and the *kanji characters* for native and Sino-Japanese words. In addition, learning these lexical classes is a part of obligatory elementary school education. On the one hand, the use of the *katakana* orthography, for example, may give rise to a sense of uniformity among recent loanwords. On the other hand, when there is a phonotactic cue that signals a specific set of loanwords from a particular language, that cue may indeed be psychologically associated with that specific sub-lexicon.

1.2 The specific hypotheses tested

The current study tested three phonotactic tendencies that can potentially cue a specific sublexicon in Japanese, which are listed in (1):

[1]For proposals regarding how such sub-lexicons may be learned, see Morita (2018), Pater (2005) and Shaw (2006).

(1) Phonotactic tendencies that can potentially cue a specific sub-lexicon

a. Singleton /p/ often appears in snack names.

b. Geminate /ɾɾ/ appears almost exclusively in loanwords from Italian.

c. Geminate /hh/ appears almost exclusively in loanwords from German.

The first observation is reported and discussed in a recent popular science book (Kawahara 2023), which grew out of the dialogue-based linguistics lecture that the author gave to elementary school children. There, one student pointed out that there are many snack names that contain singleton /p/ in Japanese (e.g. /pokkii/, /papiko/, /poihuɾu/, /poɾinkii/ and /pai-no-mi/)—the observation which we may mnemonically dub "the Pocky effect"—and she asked why. Kawahara (2023) did not offer a quantitative backup of this observation, but assuming that /p/ is indeed overrepresented in snack names in Japanese, he speculated that singleton /p/ is used to represent European—as opposed to traditional, Japanese—snack names, because, (non-post-nasal) singleton /p/ appears only in loanwords in Japanese (Ito & Mester 1995, 1999, 2008); in a sense, singleton /p/ therefore cues "foreignness" in Japanese, as the experiment by Moreton & Amano (1999) demonstrates, which the product companies may be taking an advantage of.[2] He additionally entertains the possibility that /p/ is preferred in snack names because /p/ is known to convey a sound symbolic meaning of "cuteness" in Japanese (Kawahara 2019; Kumagai 2019). Regardless of whether these conjectures offered by Kawahara (2023) are on the right track, this dialogue raised an interesting question that should be empirically tested: whether Japanese speakers indeed associate singleton /p/ with snack names, i.e. whether the observation made by the elementary school student is psychologically real or not.

As for the second hypothesis, we are not aware of any previous systematic study on the observation, but it was instead initially based on our intuitions as native speakers of Japanese. The native phonology of Japanese does not allow geminate /ɾɾ/ (Kawahara & Pangilinan 2017; Kuroda 1965; Labrune 2014), perhaps because it is a flap which is intrinsically characterized by its short duration. However, geminate /ɾɾ/ is observed in very recent loanwords, and appears mainly in names for Italian foods (e.g. /huʑiɾɾi/ 'fusulli', /ɸaɾuɸaɾɾe/ 'Farfalle' and /kitaɾɾa/ 'chitarra'). Tanaka (2017), which offers a corpus-based analysis of Italian loanwords in Japanese, shows that almost any type of geminates can appear in loanwords borrowed from Italian, except for glide geminates which Italian itself does not have.[3] This recent use of /ɾɾ/ in Japanese was

[2]This idea can be formally captured by a family of Express(x) constraints proposed by Alderete & Kochetov (2017), which requires that certain sounds be used to express a particular semantic notion; e.g. use palatal consonants and/or high front vowels to express smallness. It is not hard to imagine that a stochastic version of a constraint like Express(p)ForSnack is at work here.

[3]Geminate /ɾɾ/ can appear in any type of words, including native words, if they are created via emphatic gemination, e.g. [kaɾɾui] 'surprisingly light'. This emphatic gemination process is known to create otherwise restricted types of geminates, such as voiced obstruent geminates in native words (Kawahara 2002). When geminated, /ɾɾ/ tends to be produced with lateral-like articulations (see Morimoto 2020 for the articulatory study of geminate /ɾɾ/ in

likely prompted by the fact that Italian has geminate /ll/ and /rr/ as their phonemes (De Benedetto & De Nardis 2021).

To more objectively access our initial intuition in a quantitative manner, we have consulted the frequency list of the Balanced Corpus of Contemporary Written Japanese (BCCWJ: Maekawa et al. 2014),[4] which revealed 64 types of words containing /ɾɾ/. Among those, most, if not all, of them are borrowed from Italian, although there were 9 words whose origin was unclear after searching with Google and other tools. The only exception is /aɾuhuɾɾa/, which is likely to have come from 'al-hurra' in Arabic; in addition to this, albeit not being found in the corpus, we think that /aɾraa/ is a possible pronunciation of the Arabic word 'Allah.' However, these two seem to be the sole exceptions, perhaps because Japanese has so far borrowed more words from Italian than from Arabic. At any rate, it seems safe to conclude that /ɾɾ/ appears mostly if not exclusively appear in loanwords from Italian. Having established this connection, the experiment reported below addressed whether Japanese speakers do indeed associate /ɾɾ/ with the sub-lexicon of "Italian names," thereby also testing whether such a very specific sub-lexicon can be psychologically real.

The third hypothesis is concerned with a certain pattern found in loanwords from German—Japanese borrowed many medical, philosophical and other technical terms from German, and these borrowings can contain geminate /hh/, which generally corresponds to a word-final velar or palatal fricative in original German words. According to our intuition, such words with /hh/ are largely limited to those loanwords from German, including /bahha/ 'Bach', /mahha/ 'mach' and /ɾiibihhi/ 'Liebig', which are also characterized by copy vowel epenthesis in word-final position (Kawahara 2007). These geminates, surrounded by the identical vowels, do not occur in native words, and rarely occur in other types of loanwords.

We have searched BCCWJ for those words containing geminate /hh/, which revealed 73 types of words, many of which are based on German words. There were 9 words, such as /boitɕahha/ and /uɾatahha/, whose origin was not very clear even after searching with Google. The non-German origin loanwords containing /hh/ that we found in the corpus were /sutahhu/ 'stuff', /huɾahhu/ 'fluff' and /sunahhu/ 'snuff', which are borrowed from English, as well as /buhha/ 'Tunisian alcohol' and /ɸoɾusutahhu/ 'Falstaff, (English/Italian) opera name', but these tokens do not have identical vowels surrounding the geminate. The only robust robust exception was the name of the painter van Gogh, borrowed as /gohho/ (and another token of a compound including this word), which comes from Dutch, a West Germanic language whose phonology is similar to the closely related German. However, the Japanese form /gohho/ suggests that the adaptation

Japanese).

[4] `https://clrd.ninjal.ac.jp/bccwj/en/index.html`, consulted May 2024. Since the analysis is based on a published corpus, we are not able to make our analysis files publicly available, but they are available upon request.

channel for this particular item can be through another language (possibly German), because the first consonant of Gogh is not a stop but a fricative in Dutch. Setting this complicated example aside, at least many of the relevant tokens were of the German origin. There thus seems to be some statistical connection between geminate /hh/ and German names, which is useful for addressing the question of whether a sub-lexicon that is as specific as "German names" can be psychologically real in the minds of Japanese speakers.

2 Experiment 1

2.1 Method

2.1.1 Task

The experiment was designed to address whether Japanese speakers associate /p/, /ɾɾ/ and /hh/ with snack names, Italian names, and German names, respectively. In the current experiment, the participants were presented with one stimulus item, and were asked to judge whether that item is better suited as a name of one genre (=the target category) or the other (=the control category).

For the first hypothesis (the connection between /p/ and snack names—the Pocky effect), the participants were asked whether each name is an European snack name or a cosmetic brand name, the latter of which is used for comparison because most cosmetic names in Japan are loanwords, just like European snack names, and are usually written with the *katakana* orthography. For the second hypothesis (the connection between /ɾɾ/ and Italian names), the participants were asked wether each stimulus is a name for an Italian restaurant or a name for a French restaurant. These two types of restaurants are both common in the current Japanese community, and since French does not have geminate liquids, no borrowed words from French in Japanese contain /ɾɾ/.[5] For the third hypothesis (the connection between /hh/ and German names), the participants were asked whether a given name is a German celebrity name or an English celebrity name.

2.1.2 Stimuli

The list of the stimuli is shown in Table 1. Since we used the Buy Response function of SurveyMonkey (see below), we were limited to have 50 questions, including one question to present the consent form, one question to check qualification for participation, and another question to

[5]There are French words that contain orthographic 'rr' such as 'Pierre,' 'Pierrot,' 'marron,' and 'surréalisme', but these are borrowed with singleton /ɾ/ in Japanese, i.e., /pieeɾu/, /pieɾo/, /maɾoɴ/ and /ɕuuɾuɾeaɾisumu/. See Shinohara (2004) for extended discussion on how French words are borrowed into Japanese. There are several possible reasons why only Italian loanwords allow geminate /ɾɾ/; e.g. purely orthographic geminate 'rr' was not borrowed as geminates; phonetic realizations of 'r' differ between French and Italian; borrowing from Italian occurred more recently than borrowing from French, and Japanese speakers may have become more willing to accommodate types of geminates that their phonology otherwise does not permit.

present the instructions for the experiment. Therefore we had 8 pairs of items for the first two hypotheses (8 × 2 × 2) and 7 pairs of items for the third hypothesis (7 × 2), resulting in a total of 49 questions (3+32+14).

Table 1: The list of nonce words used in the experiment.

	Target	Control
/p/=snack	/paɾiko/	/taɾiko/
	/pasomi/	/tasomi/
	/penaɾo/	/tenaɾo/
	/posone/	/tosone/
	/pamoɾe/	/kamoɾe/
	/paɾase/	/kaɾase/
	/pesemo/	/kesemo/
	/poniɾu/	/koniɾu /
/ɾɾ/=Italian	/metoɾɾa/	/metobba/
	/makoɾɾo/	/makobbo/
	/nesoɾɾa/	/nesodda/
	/nokiɾɾe/	/nokidde/
	/temeɾɾo/	/temeddo/
	/tamiɾɾa/	/tamigga/
	/kanoɾɾi/	/kanoggi/
	/tonoɾɾe/	/tonogge/
/hh/=German	/bohho/	/boppo/
	/kuhhu/	/kuppu/
	/gehhe/	/geppe/
	/gahha/	/gatta/
	/ɾehhe/	/rette/
	/bihhi/	/bikki/
	/nahha/	/nakka/

For the first hypothesis, the target items began with /p/, whereas the control items began with either /t/ or /k/. The rest of the words was identical between the target items and the control items. We avoided using high vowels after /t/, because they cause affrication of the preceding stop (Vance 2008). For the second hypothesis, the target items included /ɾɾ/ whereas the control items contained voiced obstruent geminates, which signaled that control items were also loanwords (Ito & Mester 1995, 1999, 2008).[6] For the third hypothesis, the target items contained /hh/, whereas the control items contained voiceless stop geminates. We included geminates in the control items

[6]After the results of both Experiments 1 and 2 were all analyzed, it was pointed out by an anonymous reviewer that Italian does not have /gg/ and hence using /gg/ in the control items could have worked as a confound. For a post-hoc analysis addressing this concern, which turned out not to be a problem, see the Appendix.

for the second and third hypotheses to make sure that it is not a mere presence of any kind of geminates that cue a particular sub-lexicon.

2.1.3 Participants

The experiment was conducted online using SurveyMonkey. The participants were collected using the Buy Response function offered by SurveyMonkey. A total of 162 native speakers of Japanese (female=71; male=91), who confirmed that they have not studied either sound symbolism (related to /p/ or not) or Japanese phonetics/phonology, completed the experiment.

2.1.4 Procedure

In the instructions, the participants were told that in each trial, they are given one name and two categories and were asked to choose which category better fits that name. Example questions were thus, "given /paɾiko/, is the name better for a European snack or a cosmetic product?" and "given /metoɾɾa/, is the name better for an Italian restaurant name or a French restaurant name?" Each stimulus was presented in isolation, not in a pair, i.e. the experiment was a forced-choice task, not a two-alternative forced choice (2AFC) task. The stimuli were all written in the *katakata* orthography, which is used primarily for loanwords in the Japanese orthographic system. Although the stimuli were presented in written forms, the participants were asked to produce each form before they register their response. The order of the stimuli was randomized by SurveyMonkey.

2.1.5 Statistical analyses

For statistical analyses, we fit a Bayesian mixed effects logistic regression model, using the `brms` package (Bürkner 2017) implemented in R (R Development Core Team 1993–). For accessible introduction to Bayesian modeling, we would like to refer the readers to Franke & Roettger (2019), Kruschke (2014), Kruschke & Liddell (2018), McElreath (2020), and Vasishth et al. (2018). Simply put, Bayesian analyses use a prior distribution and the obtained data to yield a posterior distribution of a parameter that we would like to estimate.

One common way to interpret the results of Bayesian regression models is to examine the middle 95% of the posterior distribution of an estimate parameter, $\hat{\beta}$, which is known as a 95% credible interval (also known as high density interval). If its credible interval does not include zero, then that effect can be considered to be meaningful/credible. However, in Bayesian analyses, we do not need to be bound by the "credible" vs. "non-credible" dichotomy, unlike in a frequentist statistical testing with a strict "significant" vs "non-significant" dichotomy. That is, another way to interpret the results of Bayesian regression models is to calculate how many samples of the

coefficient of interest are in the expected direction in the posterior distribution. In the current paper, we present both measures to interpret the results.

The details of the current model specifications were as follows. The dependent variable was whether each item was chosen as the target category name ("snack names", "Italian names", "German names") or not. The main independent variable is the fixed factor encoding whether each item was the target (containing the phonotactic cue) or the control (not containing the phonotactic cue). We included a random intercept for items and a random intercept and slope for participants associated with that fixed factor.

For prior specifications, we used a Normal(0, 1) weakly informative prior for the intercept (Lemoine 2019) and a Cauchy prior with scale of 2.5 for the slope (Gelman et al. 2018). Four chains with 2,000 iterations were run, and the first 1,000 iterations from each chain were disregarded as warmups. All the $\hat{R}$-values were 1.00 and no divergent transitions were detected. The raw data, the R Markdown file with the R syntax as well as the posterior samples are available in an OSF repository.[7]

2.2 Results

2.2.1 Snack names

Figure 1 shows the results concerning the Pocky effect, which is the violin plot representing the distributions of "snack responses ratios" for the control items (those that begin with either /t/ or /k/) and for the target items (=those that begin with /p/). Transparent blue circles, slightly jittered to avoid overlap, represent averaged responses for each condition from each participant. Solid red circles are the grand averages in each condition and the red bars around the circles represent the bootstrap 95% confidence intervals around these averages, calculated by the `ggplot` package (Wickham 2016).

[7] `https://osf.io/97zc5/?viewonly=4efac77f1b1a4e7c80798f7efbe86160`

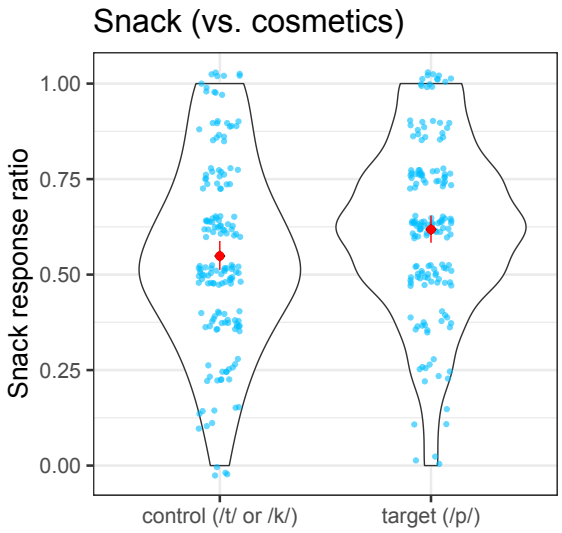

Figure 1: The comparison between the control and the target items in terms of the probabilities of being chosen as snack names, as opposed to cosmetic names.

The grand averages for the control and the target conditions were 0.55 vs. 0.62, respectively, which shows that the items beginning with /p/ were more likely to be judged as snack names than the control items.

The result of the Bayesian analysis shows that the central coefficient estimate ($\hat{\beta}$) of the difference between the control and the target is 0.33, suggesting that the target items with singleton /p/ induced more the snack responses than the control items with /t/ or /k/. The 95% credible interval of this coefficient is [-0.09, 0.75]. Although this interval includes 0, it is heavily skewed toward the positive values, and the posterior probability of this coefficient being positive ($\hat{\beta} > 0$) is 0.95.

With this result, it seems safe to conclude with a reasonable amount of confidence that Japanese speakers indeed associate names with /p/ with snack names. In other words, the psychological reality of the Pocky effect (Kawahara 2023) is corroborated by the current behavioral experiment.

2.2.2 Italian names

To assess the second hypothesis, Figure 2 shows the violin plot of the "Italian responses ratios" for the control condition and target condition, the latter of which contained /ɾɾ/. The grand average was 0.64 for the control condition vs. 0.66 for the target condition.

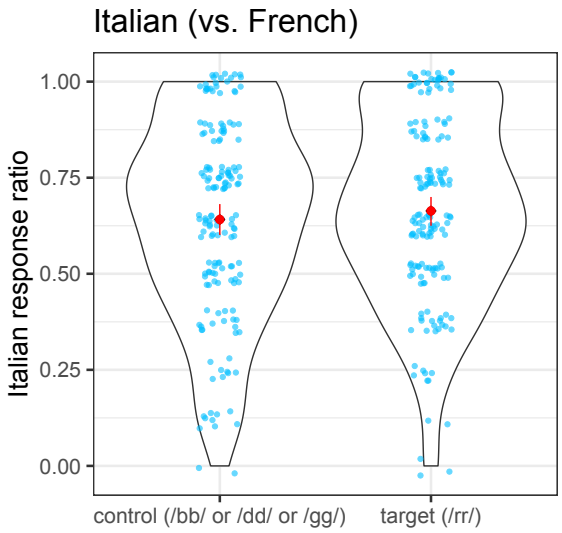

Figure 2: The comparison between the control and the target items in terms of the probabilities of being chosen as Italian restaurant names as opposed to French restaurant names.

The difference is in the expected direction—those that contain /ɾɾ/ were slightly more likely to be judged to be Italian names than those that contain voiced obstruent geminates. However, the magnitude of this difference is very small. The central coefficient estimate of this difference is 0.11, with its 95% credible interval being [-0.13, 0.35]. The posterior probability of this coefficient being positive is 0.83. The evidence for this difference thus appears to be weak or at best modest.

2.2.3 German names

Figure 3 shows the probability distribution of the German response ratios for the stimuli that were intended to address the third hypothesis. The grand average for the control items was 0.60 and the average for the target items is 0.71, suggesting that those nonce items containing /hh/, surrounded by the same vowels, were more likely to be judged as German celebrity names. The central estimate of this coefficient $\hat{\beta}$ for this difference is 0.67, with its 95% credible interval being [-0.11, 1.40]. The probability of this coefficient being positive in the posterior distribution, $p(\hat{\beta} > 0)$, is 0.95. These results suggest that Japanese speakers do indeed tend to associate /hh/ with German names, as opposed to English names.

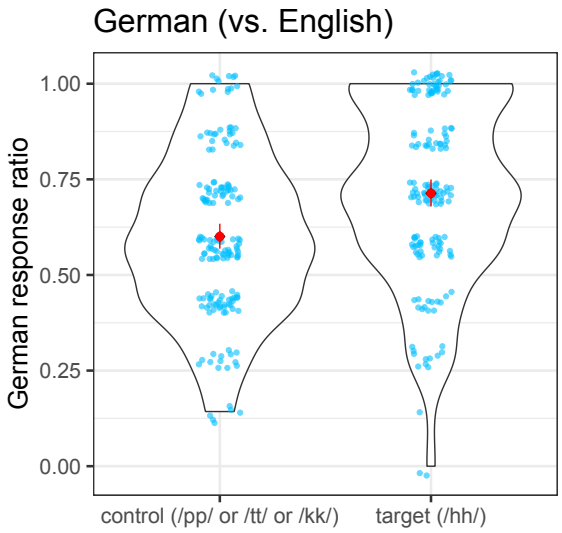

Figure 3: The comparison between the control and the target items in terms of the probabilities of being chosen as German celebrity names, as opposed to English names.

2.3 Discussion

The results of the first experiment show that we have reasonable amount of evidence to believe that Japanese speakers associate /p/ with snack names and /hh/ with German celebrity names, and we have weak or perhaps modest evidence showing that Japanese speakers associate /ɾɾ/ with Italian names. More generally speaking, the results show that Japanese speakers can associate nonce words with a particular phonotactic cue to a very specific lexical class, at least going beyond the traditional distinction between native words vs. loanwords.

The connection between /ɾɾ/ and Italian restaurant names, as revealed in this experiment, was not very robust, however. One possibility is that those who are not familiar with linguistics—which was exactly the pool of participants in this experiment—may not clearly distinguish French names and Italian names. Another possibility is that the participants may have avoided associating names with voiced obstruents with French names—if we look back the results in Figure 2, the control items (containing voiced obstruent geminates) were associated with Italian names with the probability that is much higher than chance (=0.64). And actually, with hindsight, there seems to be a good reason for them to have done so. Our post-experimental search shows, for example, that the website[8] which lists 131 popular French loanwords in Japanese, contains no instance of a French loanword containing a voiced obstruent geminate. This consideration suggests that the results of Experiment 1 may have been affected by how the control categories, rather than the

[8] `https://origamijapan.net/origami/2019/07/09/france-gairaigo/`, last access May 2024/

target categories, were judged.

For the third hypothesis also, an alternative interpretation was pointed out by our anonymous colleague—it may be the case that the participants were not so much familiar with German names, but instead they were only familiar with English names. As long as they have a general idea of how English loanwords look like, they may have realized that /hh/ is absent in such words, and may have associated those words with /hh/ with "the non-English option." This alternative, together with the possibility discussed in the preceding paragraph, point to a general methodological issue in Experiment 1: the participants may have been making judgments about the control categories (e.g. "no voiced obstruent geminates in French names" and "no /hh/ in English names") rather than the target categories.

A possibly related concern may be raised for the snack vs. cosmetic comparison as well. More concretely, it is not clear whether participants of all ages—and both genders—had the same knowledge regarding how cosmetic names generally sound like in Japanese. Familiarity with cosmetic names could have varied between speakers, and hence this variability could have worked as an additional complication, if not as a confound, in Experiment 1.

To address this general concern—that the results of Experiment 1 may have been influenced by judgments about the control categories—in the next experiment, we avoided comparing target categories and control categories, because after all, we are not directly interested in how control categories behave. Instead, we asked the participants to judge how suitable each name is for the three respective target categories.

In addition, after we ran Experiment 1, we realized that one of the control items for the third hypothesis, /geppe/, could have sounded too close to /gebbeɾusu/ 'Goebbles'—an (in)famous German historical figure—whose /bb/ can be devoiced, because of its OCP(voice) violation due to its co-occurrence with /g/ (Kawahara 2006). As a post-hoc analysis, we compared the German response ratio for /geppe/ and that of the other items, which showed the /geppe/ was judged to be a German name 81.5% of the time, while the average response for the other control items is 56.5%. The use of this particular item was thus not ideal, an issue that we also fixed in Experiment 2.[9]

3 Experiment 2

3.1 Methods

The crucial methodological differences between Experiment 1 and Experiment 2 were as follows. Experiment 2 only tested the three target categories ("snack names", "Italian names" and "German names") instead of comparing them to the control categories, because the results of Experiment

[9]We refrain from running a new statistical analysis without /geppe/, to avoid HARKing (Hypothesizing After the Results are Known) (Kerr 1998).

1 may have been influenced by how the control categories were judged. Concretely, in this next experiment, we presented each stimulus and asked how suitable each name is for the three respective target categories, using a 4-point Likert scale, with the following labels: 1="not at all suitable", 2="not so suitable", 3="suitable" and 4="very suitable." Example questions are thus, "How suitable is /paɾiko/ as a snack name?" and "How suitable is /bohho/ as a German celebrity name?" We used the same set of stimuli as Experiment 1 (see Table 1), except that for the reason discussed above, we replaced /gehhe/ and /geppe/ with /gihhi/ and /gippi/, respectively. As with Experiment 1, the participants were asked to read each stimulus before registering their responses.

For this experiment, we used a snow-balling sampling method primarily using the first author's X account (formerly Twitter). The data from 162 native speakers of Japanese were collected (gender information was not asked for this experiment). Since the responses were obtained using a Likert scale, we used an *ordinal* logistic regression. Other details of the statistical analysis were identical to that of Experiment 1. The files used for the statistical analyses for this experiment are also made available at the above-mentioned OSF repository.

3.2 Results

3.2.1 Snack names

Figure 4 is a violin plot showing how suitable the control items (with /t/ or /k/) and the target items (with /p/) were judged as snack names, i.e. a test of the Pocky effect. We observe that the latter items were judged to be more suitable (the grand averages were 2.53 vs. 2.86, respectively). The central estimate of this coefficient ($\hat{\beta}$) for this difference is 0.86 with its 95% credible interval being [-0.44, 2.19]. The posterior probability of this coefficient being positive ($\hat{\beta} > 0$) is 0.91.

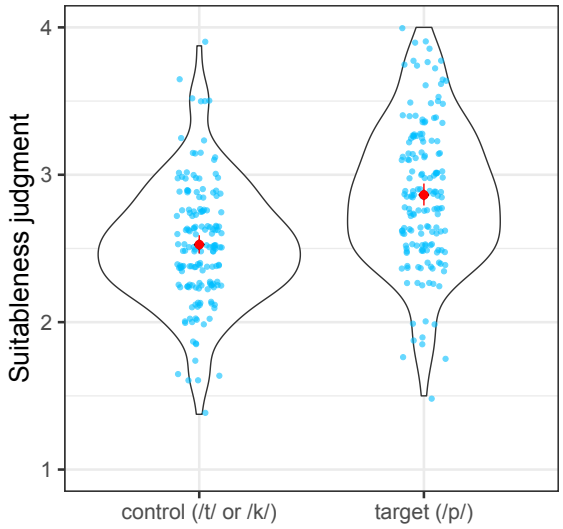

Figure 4: The comparison between the control and the target items in terms of how suitable they were judged as snack names.

3.2.2 Italian names

Figure 5 shows how suitable the control items (with a voiced obstruent geminate) and the target items (with geminate /ɾɾ/) were judged as the names of Italian restaurants. Compared to the results of Experiment 1, we observe a rather substantial difference between the two conditions (2.22 vs. 2.94). The central coefficient estimate for this difference was 2.12, with its 95% credible interval being [1.56, 2.69]. All the posterior samples were positive ($\hat{\beta} > 0 = 1$), suggesting that the connection between /ɾɾ/ and Italian names is very robust. This result is consistent with the post-hoc speculation we made about Experiment 1; namely, the participants may have avoided associating names with a voiced obstruent geminate with French restaurant names, hence shrinking the difference between the control items with a voiced geminate and the target items with /ɾɾ/ in Experiment 1.[10]

[10] An anonymous reviewer asked if voiced geminates may have been judged to be not very suitable for Italian names in Experiment 2. With the average rating being 2.22, this may have been the case, although we note that in the German condition, the average rating for the control items was comparable (i.e. 2.35) and also that in Experiment 1, voiced obstruent geminates were not particularly disfavored as Italian names. If voiced geminates were indeed judged to be unsuitable for Italian names, however, it raises a question regarding whether the difference that we observe in Figure 5 is truly due to the effects of /ɾɾ/. A follow-up experiment with different control items, perhaps with a voiceless obstruent geminate, is necessary to address this concern. See also the Appendix.

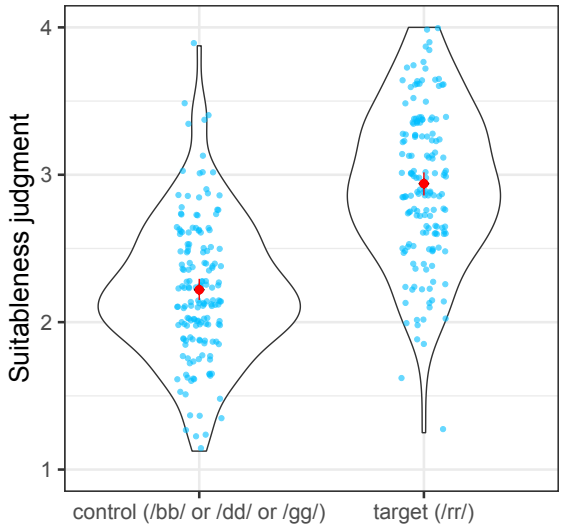

Figure 5: The comparison between the control and the target items in terms of how suitable they were judged as Italian restaurant names.

3.2.3 German names

Figure 6 shows the difference between the control items and the target items in terms of how suitable they were judged as German names. Those target items with /hh/ were judged to be more suitable than the control items with voiceless stop geminates (2.93 vs. 2.35); the central coefficient estimate is 1.62, with its 95% credible interval being [0.62, 2.57]. More than 99% of the posterior samples were positive, showing that this association is very robust.

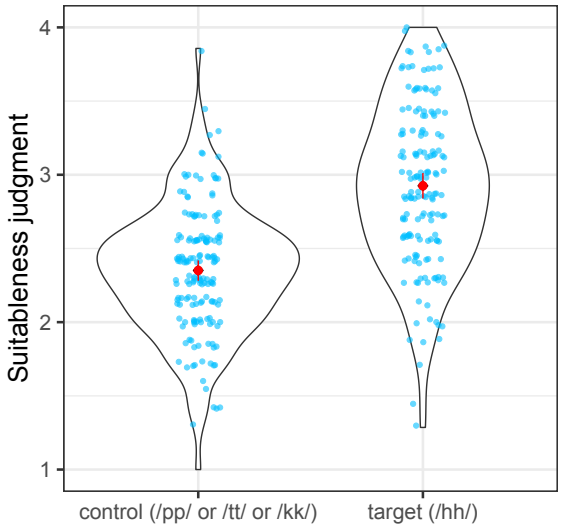

Figure 6: The comparison between the control and the target items in terms of how suitable they were judged as German names.

3.3 Discussion

The results were all in the expected direction and were generally clearer in Experiment 2 than in Experiment 1. Overall, the results suggest that Japanese speakers can infer, based a particular phonotactic cue, a very specific subpart of the lexicon, that is something as specific as "snack names", "Italian names" and "German names". In this sense, the traditional lexical stratification into native, Sino-Japanese, loanwords and mimetics does not suffice to explain the current experimental results.

4 General discussion

We started with a general question concerning how the lexicon can be organized into smaller sub-lexicons in a natural language. Phonotactic restrictions do not hold uniformly across the lexicon of a whole language and the lexicon thus seems to be organized into sub-lexicons—then, how fine-grained can this sub-division be? The classic, and conservative, view is that such division should be very general, perhaps with independent etymological/orthographic motivations (Chomsky & Halle 1968; Ito & Mester 1995, 1999, 2008). The alternative view is that the sub-lexicons can be very specific, as specific as "a set of roots that a particular affix is attached to" (Becker & Gouskova 2016; Gouskova et al. 2015; Inkelas et al. 1996; Inkelas & Zoll 2007; Pater 2005).

To address this general question, the paper first pointed out three cases in which particu-

lar phonotactic tendencies can cue a very specific word-class in Japanese, two of which were quantitatively examined by a corpus-study. We then moved on to two experiments, which have demonstrated that Japanese speakers can infer specific lexical classes such as "snack names" and "German names" based on phonotactic cues, which is compatible with the second view.

The current study opens up several opportunities for future studies. For instance, it would not be surprising if Japanese speakers differ in terms of how familiar they are with the Italian or German cultures, and if so, they have different degrees of knowledge about the relevant sub-lexicons. Other factors, such as socio-economic statuses and educational levels may influence how these sub-lexicons are structured. Generally speaking, then, future studies should explore the inter-speaker variability that may exist regarding the sort of sub-lexicons that are identified in the current experiment.

Moreover, the current study focused on three specific cases of phonotactic cues that are associated with a specific sub-lexicon, but it is not hard to imagine that there can be other similar cases in Japanese and other languages. For example, it has been pointed out that back vowels are overrepresented in ice cream names in English (Jurafsky 2014), and therefore, it would be interesting to explore whether English speakers can use this cue to identify an ice cream name. It is also not hard to imagine that among the loanwords in English, forms with only light syllables may be overrepresented in those that are borrowed from Japanese, such as *Toshiba*, *Pikachu* and *Yamaha*. It would be interesting to test whether English speakers would tend to identify new nonce words with only light syllables as those borrowed specifically from Japanese, while also examining the degree to which the familiarity with the Japanese culture influences the strength of such associations. In short, more experiments are warranted to address just how fine-grained sub-lexicons in natural languages can be, and how these sub-lexical structures may or may not vary between speakers of the same speech community.

Appendix

After all the analyses were completed, it was pointed out by an anonymous reviewer that Italian does not have /gg/ in the first place. While both /bb/ and /dd/ are attested in Italian loanwords in Japanese, /gg/ is thus not (Tanaka 2017). This raises the question regarding whether the items containing /gg/ were disfavored as Italian names in our experiments. Noting that these are post-hoc analyses, we compared the responses for the items containing /bb/, /dd/ and /gg/, separately. In Experiment 1, the average Italian responses were: /bb/=0.62, /dd/=0.63, /gg/=0.67, which actually indicates that the items with /gg/ were more likely to be associated with Italian names than those with /bb/ or /dd/. The average ratings in Experiment 2 were comparable across the three types of voiced obstruent geminates: /bb/=2.22, /dd/=2.21, /gg/=2.23. These results imply that

linguistically naive native speakers of Japanese are probably unaware that Italian loanwords do not contain /gg/. These are interesting results in and of themselves, but we reiterate that these are post-hoc analyses and do not make any conclusive statements. These results may imply, however, that it is possible that not all phonotactic aspects of a specific set of loanwords are learned.

Funding

This project is supported by JSPS grants #22K00559 to the first author and #19K13164 to the second author.

Conflicts of interest

We declare no conflicts of interest.

Availability of data and material

The data are available at
`https://osf.io/ym79p/?viewonly=ce17de5a39834ae397c44a19e74db082`

Code availability

The code is also available at
`https://osf.io/ym79p/?viewonly=ce17de5a39834ae397c44a19e74db082`

Authors' contributions

Both authors contributed to the conception and execution of the experiments. The first author wrote the manuscript and the second author revised it. Both authors worked on the revision. The statistical analysis was primarily conducted by the first author. The second author checked the details.

Ethics approval

The current experiments were conducted with an approval from the first author's institute.

Consent to participate

The participants read the written consent form before participating in the experiments.

Consent for publication

Both authors approve that the current manuscript be evaluated for publication in the journal.

References

Alderete, John & Alexei Kochetov. 2017. Integrating sound symbolism with core grammar: The case of expressive palatalization. *Language* 93. 731–766.

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20(1). 1–42.

Becker, Michael & Maria Gouskova. 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* 47(3). 391–425.

Blake, Allen & Michael Becker. 2015. Learning alternations from surface forms with sublexical phonology. Ms. lingbuzz/002503.

Bürkner, Paul-Christian. 2017. brms: An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software* 80(1). 1–28.

Chomsky, Noam & Moris Halle. 1968. *The sound pattern of English.* New York: Harper and Row.

De Benedetto, Maria Gabriella & Luca De Nardis. 2021. Consonant gemination in Italian: The nasal and liquid case. *Speech Communication* 133. 62–80.

Franke, Michael & Timo B. Roettger. 2019. Bayesian regression modeling (for factorial designs): A tutorial. Ms. `https://doi.org/10.31234/osf.io/cdxv3`.

Gelbart, Ben. 2005. *Perception of foreignness*: University of. Massachusetts, Amherst Doctoral dissertation.

Gelbart, Ben & Shigeto Kawahara. 2007. Lexical cues to foreignness in Japanese. In *Proceedings of formal approaches to Japanese linguistics 4 [MIT working papers in linguistics 55]*, 49–60. Cambridge: MITWPL.

Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau & Yu-Sung Su. 2018. A weakly informative default prior distribution for logistic and other regression models. *Annual Applied Statistics* 2(4). 1360–1383.

Gouskova, Maria & Suzy Ahn. 2024. Sublexical phonotactics and English comparatives. *University of Massachusetts Occasional Papers in Linguistics 50th Anniversary Volume* .

Gouskova, Maria, Sofya Kasyanenko & Luiza Newlin-Lukowicz. 2015. Selectional restrictions as phonotactics over sublexicons. *Lingua* 41–81.

Inkelas, Sharon, Orhan Orgun & Cheryl Zoll. 1996. Subregularities as cogrammars: the theoretical status of nonproductive patterns and exceptions in grammars. In *Proceedings of the nels 26*, Ms., University of California, Berkeley.

Inkelas, Sharon, Orhan Orgun & Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In Iggy Roca (ed.), *Derivations and constraints in phonology*, 393–418. Oxford: Oxford University Press.

Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45. 133–172.
Ito, Junko & Armin Mester. 1995. Japanese phonology. In John Goldsmith (ed.), *The handbook of phonological theory*, 817–838. Oxford: Blackwell.
Ito, Junko & Armin Mester. 1996. Stem and word in Sino-japanese. In Takashi Otake & Anne Cutler (eds.), *Phonological structure and language processing: Cross-linguistic studies*, 13–44. Berlin: Mouton de Gruyter.
Ito, Junko & Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 62–100. Oxford: Blackwell.
Ito, Junko & Armin Mester. 2008. Lexical classes in phonology. In Shigeru Miyagawa & Mamoru Saito (eds.), *The Oxford handbook of Japanese linguistics*, 84–106. Oxford: Oxford University Press.
Jurafsky, Dan. 2014. *The language of food: A linguist reads the menu.* New York: W. W. Norton & Company.
Kawahara, Shigeto. 2002. Similarity among variants: Output-variant correspondence. BA thesis, International Christian University, Tokyo Japan.
Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82(3). 536–574.
Kawahara, Shigeto. 2007. Copying and spreading in phonologicla theory: Evidence from echo epenthesis. In Leah Bateman, Michael O'Keefe, Ehren Reilly & Adam Werle (eds.), *University of Massachusetts occasional papers in linguistics 32: Papers in Optimality Theory iii*, 111–143. Amherst: GLSA.
Kawahara, Shigeto. 2019. What's in a PreCure name? *ICU Working Papers in Linguistics 7: Festschrift for Professor Tomoyuki Yoshida on his 60th birthday* 15–22.
Kawahara, Shigeto. 2023. *Why snack names often contain [p] in Japanese: A linguist interacts with elementary school students [in Japanese].* Discover 21.
Kawahara, Shigeto & Melanie Pangilinan. 2017. Spectral continuity, amplitude changes, and perception of length contrasts. In Haruo Kubozono (ed.), *Aspects of geminate consonants*, 13–33. Oxford: Oxford University Press.
Kerr, N.L. 1998. HARKing: Hypothesizing after the results are known. *Personality and Psychology Review* 2(3). 196–217.
Kruschke, John K. 2014. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan.* Waltham: Academic Press.
Kruschke, John K. & Torrin M. Liddell. 2018. The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychological Bulletin and Review* 25. 178–206.
Kumagai, Gakuji. 2019. A sound-symbolic alternation to express cuteness and the orthographic Lyman's Law in Japanese. *Journal of Japanese Linguistics* 35(1). 39–74.
Kuroda, S.-Y. 1965. *Generative grammatical studies in the Japanese language*: MIT Doctoral dissertation.
Labrune, Laurence. 2014. The phonology of Japanese /r/: A panchronic account. *Journal of East Asian Linguistics* 23(1). 1–25.
Lemoine, Nathan P. 2019. Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos* 128. 912–928.

Maekawa, Kikuo, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka & Yasuharu Den. 2014. Balanced corpus of contemporary written Japanese. *Language Resources and Evaluation* 48(2). 345–371.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan, 2nd edition*. London: Taylor & Francis Ltd.

Moreton, Elliot & Shigeaki Amano. 1999. Phonotactics in the perception of Japanese vowel length: Evidence for long distance dependencies. *Proceedings of the 6th European Conference on Speech Communication and Technology* .

Morimoto, Maho. 2020. *Geminated liquids in Japanese: A production study*: University of California, Santa Cruz Doctoral dissertation.

Morita, Takashi. 2018. *Unsupervised learning of lexical subclasses from phonotactics*: MIT Doctoral dissertation.

Orgun, C. Orhan. 1996. *Sign-based morphology and phonology: With special attention to Optimality Theory*: University of California, Berkeley Doctoral dissertation.

Pater, Joe. 2005. Learning a stratified grammar. In Alejna Brugos, Manuella R. Clark-Cotton & Seungwan Ha (eds.), *Proceedings of the 29th boston university conference on language development*, 482–492. Somerville: Cascadilla.

R Development Core Team. 1993–. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria.

Sande, Hannah. 2020. Cophonologies by Ph(r)ase. *Natural Language and Linguistic Theory* 38. 1211–1261.

Shaw, Jason. 2006. Learning a stratified lexicon. *Proceedings of NELS* 36. 519–530.

Shinohara, Shigeko. 2004. Emergence of universal grammar in foreign word adaptations. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in phonological acquisition*, 292–320. Cambridge: Cambridge University Press.

Tanaka, Shin'ichi. 2017. The relation between L2 perception and L1 phonology in Japanese: An analysis of geminates in loanwords from Italian. In Haruo Kubozono (ed.), *Aspects of geminate consonants*, 321–339. Oxford: Oxford University Press.

Vance, Timothy. 2008. *The sounds of Japanese*. Cambridge: Cambridge University Press.

Vasishth, Shravan, Bruno Nicenboim, Mary Beckman, Fangfang Li & Eun Jong Kong. 2018. Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics* 71. 147–161.

Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.