

# AS HIPÓTESES SOBRE O PROTO-NIPO-COREANO: DESAFIOS NAS CORRESPONDÊNCIAS SONORAS DE PRONOMES E NUMERAIS THE HYPOTHESES ABOUT PROTO-JAPANESE- KOREAN: CHALLENGES IN SOUND CORRESPONDENCES OF PRONOUNS AND NUMERALS

*Marcus TANAKA DE LIRA<sup>1</sup>*

**Resumo:** Desde Arai (1717), diversos pesquisadores têm levantado a possibilidade de que a família Japônica e a família Coreânica possam compartilhar uma origem comum. Nestes mais de trezentos anos de pesquisa, houve avanços com o uso de técnicas do método comparativo em linguística histórica, como Martin (1966), porém um argumento mais robusto para a confirmação de uma origem comum a essas línguas, uma protolíngua cujos reflexos possam ser previstos através de regras de correspondência fonológica, permanece como um desafio. Como exemplo de como uma protolíngua é formulada e quais passos ainda são necessários para fortalecer a hipótese de um ancestral comum entre a língua japonesa e a língua coreana, alguns pronomes interrogativos e numerais cardinais baixos, de 1 a 10, foram escolhidos sistematicamente. Apesar de haver pelo menos duas correspondências sonoras promissoras nos pronomes, não parece haver cognatos nos sistemas numéricos, o que sugeriria ou um conjunto de inovações completas, eliminando quaisquer cognatos dentro dessas palavras, ou uma origem para ambas as línguas que pode não ser mais detectada com base nas ferramentas atualmente disponíveis na linguística histórica.

**Palavras-chave:** Família Japônica; Família Coreânica; Linguística Histórica; Proto-Nipo-Coreano; Nordeste Asiático

**Abstract:** Ever since Arai (1717), several researchers have put forth the possibility that the Japonic family and the Koreanic family may share a common origin. In these three hundred

---

1 Instituição: Universidade de Brasília; ORCID: <https://orcid.org/0000-0003-1750-8727>;  
e-mail: [tanakadelira@unb.br](mailto:tanakadelira@unb.br)

plus years of research, there have been breakthroughs with the use of techniques from the comparative method in historical linguistics, such as Martin (1966) , but a stronger argument for the confirmation of a common origin to these languages, a protolanguage whose reflexes can be predicted through sound correspondence rules, remains a challenge. As a sample of how a protolanguage is formulated and which steps remain necessary to strengthen the hypothesis of a common ancestor between the Japanese language and the Korean language, some interrogative pronouns and low cardinal numerals, from 1 to 10, have been systematically chosen. Despite having at least two promising sound correspondences in the pronouns, there do not seem to be any cognates in the numeral system, which would either suggest a wholesale set of innovations, eliminating any cognates within these words, or an origin for both languages that may no longer be detected based on the tools currently at the disposal of historical linguistics.”.

**Keywords:** Japonic Family; Koreanic Family; Historical Linguistics; Proto-Japanese-Korean; Northeast Asia

Qual é a relação entre a língua japonesa e as outras línguas do mundo? Embora essa pergunta pareça razoavelmente simples, a língua japonesa apresenta uma situação bastante peculiar entre as vinte línguas mais faladas no mundo, por ser a única para a qual não temos uma resposta bem aceita até os dias de hoje (SHIBATANI, 1990, p. 94; IWASAKI, 2006, p. 3-4). Apesar de existir uma vasta e extensa literatura em língua inglesa, que desempenha atualmente o papel de língua franca na ciência, por exemplo, não há muita informação disponível em português que permita um entendimento mais amplo do estado da arte contemporâneo sobre a origem da língua japonesa. A escassez de material, de certa forma, reflete a falta de consenso sobre o quanto de fato se sabe (ou não) sobre os estágios anteriores ao chamado japonês clássico. Como era a situação linguística no arquipélago nipônico antes da introdução da escrita, e qual seria a sua relação com a região do nordeste asiático como um todo?

Essa situação de desconhecimento sobre as origens e relações da língua japonesa é extremamente incomum entre as línguas mais faladas do mundo. Isso não significa que as línguas isoladas, que têm relações com outras famílias linguísticas, sejam completamente desconhecidas. No próprio Japão, há o caso da língua Ainu, que hoje está praticamente restrita à ilha de Hokkaido, e que se enquadra exatamente nesse status (DOUGHERTY, 2018, p. 100; BUGAEVA, 2012, p. 461). No entanto, essas línguas isoladas representam uma minoria atualmente, com apenas 159 línguas desse tipo em um universo de aproximadamente 7 mil línguas, e elas não fazem parte de nenhuma das 248 famílias linguísticas já identificadas (CAMPBELL, 2018, p. 1; MORAVCSIK, 2013, p. 17)<sup>2</sup>. Vale ressaltar também que a grande maioria dessas línguas isoladas são minoritárias, enquanto as línguas nacionais geralmente têm suas histórias mais bem compreendidas.

---

2 Não se consideram tradicionalmente como isoladas línguas pertencentes a famílias cujos outros membros estão atualmente extintos, mas cuja existência foi documentada, como a língua Ket na família Ienisseiana (GEORG, 2018, p. 140).

Nós sabemos que o português, a língua na qual este trabalho está escrito, deriva do latim vulgar e tem como línguas irmãs o espanhol, o italiano e o romeno (algumas das chamadas línguas românicas). Por sua vez, é sabido que o latim descende do que chamamos de Proto-Indo-Europeu, e suas línguas irmãs vão desde o grego antigo até o sânscrito (CLACKSON, 2007, p. 1). A definição de protolíngua, como ancestral comum de diferentes línguas, e a explicação de que como ela pode ser identificada e reconstruída será aprofundada na seção metodológica.

Por ora, podemos adiantar que é conhecido que o português, assim como a maioria das línguas que se estendem desde a Islândia e Portugal, no Noroeste, até Bangladesh e Nepal, no Sudeste, pertence ao tronco Indo-Europeu, contando hoje com aproximadamente três bilhões de falantes (KAPOVIĆ, 2017, p. 1). O alcance das principais famílias linguísticas é tão amplo que apenas cerca de 1 em cada 8 pessoas no mundo, aproximadamente, fala uma língua que não pertence a um dos cinco principais troncos linguísticos: Indo-Europeu, Sino-Tibetano, Niger-Congo, Afro-Asiático e Austronésio (EBERHARD, SIMONS e FENNIG, 2022).

Por um lado, existe um programa de pesquisa bastante produtivo que agrupa as línguas do mundo de forma a entender o desenvolvimento e a dispersão delas, conhecido como linguística histórica (CAMPBELL, 1998; CROWLEY e BOWERN, 2010; BOWERN e EVANS, 2015). Esse programa tem tido sucesso em classificar a grande maioria das línguas contemporâneas. Por outro lado, a posição da língua japonesa na(s) grande(s) árvore(s) de descrição da história das línguas naturais ainda é amplamente debatida. Portanto, é necessário adquirir um conhecimento mais aprofundado sobre o assunto.

A situação se torna ainda mais peculiar quando consideramos que entre as línguas com mais de 50 milhões de falantes, há apenas duas cujas origens e relações ainda são pouco conhecidas: a língua japonesa, foco principal neste trabalho, e a língua coreana (WHITMAN, 2012, p. 24), que enfrenta desafios similares. Devido à proximidade geográfica e tipológica, além desse mistério comum, não é surpreendente que haja comparações entre as duas línguas há pelo menos 300 anos (LEE e RAMSEY, 2011, p. 26; SHIBATANI, 1990, p. 94). Isso levou à criação de uma extensa literatura com reconstruções de uma língua ancestral comum às duas línguas, principalmente a partir de Martin (1966). Embora possa ter ocorrido um declínio no interesse a partir da década seguinte no Japão, com a visão de que a língua japonesa pode ter se formado em um período anterior ao que pode ser analisado pelas ferramentas atuais (HASEGAWA, 2015, p. 7), com o papel central do questionamento dentro da linguística japonesa e da identidade japonesa ressurgindo quando se encontram novas pistas sobre a origem e difusão da língua japonesa (LEE e HASEGAWA, 2011; WADE, 2011). No exterior, no entanto, a origem e a classificação da língua japonesa tem sido temas perenes na linguística histórica (ROBBEETS, 2008, p. 337), o que reforça a importância da pesquisa apresentada aqui.

Em síntese, as próximas páginas têm o propósito principal de proporcionar um panorama breve do atual estado das pesquisas sobre a origem da língua japonesa e, particularmente, das especulações acerca da sua possível relação com a língua coreana. Além disso, busca-se apresentar os desafios que impedem de afirmar não apenas que essas línguas possuem parentesco, mas também, como frequentemente discutido na literatura, que ambas fazem parte de um tronco linguístico mais amplo. Para isso, abordaremos de maneira mais específica a dificuldade em realizar uma tarefa que seria simples se a relação entre as duas famílias linguísticas fosse mais clara: estabelecer regras de mudanças sonoras regulares entre as línguas, por meio da análise de partes dos sistemas pronominal e numérico das famílias. Essa análise deveria permitir, de alguma forma, a reconstrução de um ancestral comum. Conforme veremos, tal empreendimento se mostra altamente problemático.

Para uma melhor compreensão da discussão e dos desafios na reconstrução de um possível ancestral das duas línguas, será apresentada primeiramente uma revisão bibliográfica na próxima seção. Em seguida, será oferecida uma breve explicação sobre linguística histórica e sobre como funciona o método comparativo, que será amplamente utilizado aqui para a comparação das línguas e a reconstrução de um ancestral comum. Posteriormente, serão apresentados os dados das línguas consideradas neste estudo e, conseqüentemente, será realizada a análise desses dados. Na conclusão, será oferecido um resumo da discussão presente neste trabalho, bem como possíveis direções para futuras pesquisas sobre o tema.

## 1. Revisão Bibliográfica

Toda língua está sujeita a mudanças e, como tal, a língua japonesa também não tem estado imune às transformações ao longo do tempo. Para ilustrar esse ponto, podemos utilizar a seguinte frase do Man'yōshū, que demonstra o quanto a língua japonesa evoluiu ao longo de pouco mais de uma dúzia de séculos:

I. 咲くべくなりにてあらずや<sup>3</sup> (5/0829)

Saku                      be-ku                      nari-nite                      ara-zu                      ya  
florescer. CONCL    NEC-ACOP.INF    tornar-PERF.GER  
existir-NEG.FINQ

“Não deveriam ter começado a florescer?” (FRELLESVIG, 2010, p. 71)

---

3 Originalmente escrito em caracteres chineses como “烏梅能波奈 佐企弓知理奈波 佐久良<婆那> 都伎弓佐久倍久 奈利尔弓阿良受也”. Uma vez que se trata de uma análise da língua falada, e não do meio utilizado para sua transcrição, a história da escrita da língua japonesa não será aqui analisada ou comentada, com a utilização da escrita contemporânea acima feita de forma anacrônica apenas para conveniência dos leitores.

Ainda que o exemplo acima possa ser compreendido pelos falantes do japonês contemporâneo, uma vez que as raízes verbais escolhidas no trecho são similares às atuais, muitas das estruturas gramaticais presentes no exemplo acima, transcritas de um texto do século VII, não existem mais na língua contemporânea ou mudaram significativamente. Não é difícil visualizar que, após milênios, as mudanças acumulam-se ainda mais na história de uma língua.

E toda língua muda de forma irregular. Naturalmente, há fatores como classe social (KERSWILL, 2007), gênero (COATES, 2007) e idade (LLAMAS, 2007) mas também há diferenças regionais, que ao longo do tempo podem ser consideradas como diferentes dialetos ou, dependendo dos critérios de classificação escolhidos (geralmente, a falta de inteligibilidade mútua), até mesmo diferentes línguas (WARDHAUGH, 2006, p. 25-31).

Uma aparente vantagem presente no caso das línguas japonesa e coreana é o fato de que, assim como mostrado no exemplo retirado do Man'yōshū, há um longo histórico de registros de formas antigas dessas línguas, datando do século VII e XI, respectivamente (MIYAKE, 2020, p. 9), fornecendo aproximadamente um milênio de história. Esses documentos, tratados com devido cuidado, podem nos dar um vislumbre de como eram as línguas em um período mais próximo de sua separação, caso realmente sejam oriundas de um ancestral comum. Se for o caso, cálculos de análise bayesiana, datam o início de uma suposta divergência entre as famílias coreânica e japônica em torno de 4500 anos atrás, supondo que a separação dessas das outras línguas consideradas altaicas tenha ocorrido há 7000 anos (ROBBEETS, 2020, p. 38). Entretanto, apesar de termos uma parte razoável dessas línguas registrada em documentos históricos, algo raramente disponível para as línguas do mundo, a história começa muito depois de uma possível separação, tornando o trabalho de reconstrução ainda mais desafiador.

Para entender como essa história se desenrolou no nordeste asiático, é necessário compreender melhor o que de fato é uma língua e como as famílias japônica e coreana são constituídas.

## **1.1 Dialetos, Línguas e a Família Japônica**

Como mencionado na subseção anterior, as línguas tendem a ser faladas de maneiras diferentes em vastos territórios. E para entender como isso ocorre, é preciso definir alguns termos e compreender como a língua japonesa se relaciona com eles.

Cada falante dentro de um grupo determinado tem seu modo próprio de falar, com características específicas, chamado de idioleto (BARBER e RAMIREZ, 2021). Esses idioletos individuais formam redes maiores de idioletos mutuamente inteligíveis, chamados de dialetos ou línguas (FRANÇOIS, 2015, p. 168) Há várias dificuldades na diferenciação entre dialetos e línguas, às vezes de natureza linguística, às vezes de natureza política, mas, seguindo a definição de sistemas mutuamente inteligíveis, serão considerados aqui como dialetos aqueles em que

a comunicação entre os falantes é possível. Caso contrário, considerar-se-á que os falantes utilizam diferentes línguas. Com base nessa definição, é necessário analisar como ela se aplica à língua japonesa.

As diferentes respostas começam já aqui. Não é incomum que a língua japonesa seja vista como uma língua isolada, ou seja, uma língua sem parentes conhecidos (CAMPBELL, 1998, p. 165). O problema é que, com base em fatores linguísticos, tal classificação não é inteiramente adequada.

Antigamente, nos casos em que a língua japonesa era considerada isolada, como em Shibatani (1990, p. 189-196), e aqui mencionada apenas para fins de ilustração, isso se devia à visão desses autores de que as línguas das ilhas de Okinawa eram consideradas dialetos, definição essa oriunda de sua ligação política e afinidades gramaticais com a língua japonesa. Entretanto, o reconhecimento de que esses ditos “dialetos” não são mutuamente inteligíveis com a língua japonesa é hoje ponto pacífico nos estudos sobre a família japônica, considerando-se, portanto, de que se tratam de línguas independentes (GRIMES, 1992, p. 648-650; SHIMOJI, 2012, p. 351; LAWRENCE, 2012, p. 381).

É importante observar que o fato de a língua japonesa não ser considerada uma língua isolada não se deve à impossibilidade deste tipo de língua, ou ao fato de que elas sejam raras. Muito pelo contrário, tal status não é exatamente uma situação excepcional no nordeste asiático, um exemplo clássico sendo a já mencionada língua Ainu, cujo grupo étnico no passado se estendia do Japão até o sul da península de Kamchatka e o longo das Ilhas Curilas (JANHUNEN, 2022, p. 58), assim como boa parte do agrupamento de línguas ao norte do Japão conhecidas como línguas paleossiberianas (COMRIE, 1981, p. 239).

### **1.1.2 Definindo a Família Japônica**

Como mencionado acima, o que impede a classificação dos falares de Okinawa como “dialetos” da língua japonesa, e a sua consequente classificação como uma língua isolada é o fato de que, devido à falta de inteligibilidade mútua, também é possível considerar que as diferentes línguas no arquipélago Ryūkyū se tratam de línguas aparentadas, e não de variações de uma mesma língua (CAMPBELL, 2018, p. 3). Essas línguas, como o okinawano, por sua vez formariam junto com o japonês a família japônica, ou Japonesa-Ryukyana (IWASAKI, 2006, p. 2-3; WHITMAN, 2012, p. 24), dentro da qual as línguas de Okinawa teriam uma relação distante com os dialetos do sul de Kyūshū, sendo sua língua-mãe, o Kyūshū-Ryūkyū, irmã – e não descendente – do japonês antigo (DE BOER, 2020, p. 55-57). A estrutura mais frequentemente aceita da família japônica ficaria, portanto, da seguinte forma (HEINRICH, MIYARA e SHIMOJI, 2015, p. 1; SHIMOJI, 2012, p. 351; LAWRENCE, 2012, p. 381; DE BOER, 2020, p. 46):

- Família Japônica
  - **Línguas Japonesas:**
    - Nihongo (Língua Japonesa)
    - Hachijô
- Línguas Ryukyuanas
  - **Línguas Ryukyuanas Setentrionais**
    - Amami
    - Okinawano;
  - **Línguas Ryukyuanas Meridionais**
    - Miyako
    - Yaeyama
    - Yonaguni / Dunan

Geralmente as línguas do arquipélago de Ryûkyû são vistas como fazendo parte de um contínuo e se encontram atualmente em diferentes graus de riscos de extinção (SHIMOJI, 2012, p. 351), motivo pelo qual nem mesmo a estrutura tende a ser unânime em descrições da família japônica.

Como um adendo, é necessário enfatizar que a língua japonesa não é falada de forma homogênea em todo o país. Há uma rica variedade dialetal que se encaixa em dois grupos: dialetos japoneses orientais e ocidentais (MATSUMORI e ONISHI, 2012; IWASAKI, 2006, p. 1; SHIBATANI, 1990, p. 187, 197; TANAKA DE LIRA e TANAKA DE LIRA, 2016). Estes são delimitados por diferentes isoglossas que atravessam a região de Chûbu na ilha de Honshu, convergindo aproximadamente na divisa das atuais prefeituras<sup>4</sup> de Gifu e Nagano. Da mesma forma que há essa variação dialetal na língua japonesa contemporânea, é razoável supor que etapas anteriores também deveriam ter algum grau de variação interna, já que essa diversificação linguística é uma etapa essencial no processo. Mas quais línguas fora do arquipélago japonês podem compartilhar ancestrais comuns por meio desse processo de diversificação? A candidata mais próxima, tanto geográfica quanto tipologicamente, seria a língua coreana, o que levanta os problemas discutidos na próxima subseção.

## 1.2 A Família Coreânica e Relações com a Família Japônica

4 Em respeito à terminologia adotada tanto pelo governo japonês em traduções para línguas indo-europeias, como o inglês, quanto à mantida pelo governo americano durante a ocupação japonesa de 1945 a 1952, as divisões administrativas do Japão, representadas por 都 (to), 道 (dô), 府 (fu) e 県 (ken) serão aqui chamadas coletivamente de prefeituras (TÔKEIKYOKU, 2023; JAPANESE GAZETEER, 1946; HENSHALL, 2004, p. 155; COUNCIL OF LOCAL AUTHORITIES FOR INTERNATIONAL RELATIONS, 2004), cognato do francês “prefecture”, com o entendimento de ter sido a França um dos países cuja configuração organizacional foi adotada pelo governo Meiji (WESTNEY, 1998, p. 130-132) reservando o termo províncias para as divisões relevantes anteriores à Restauração Meiji.



A península coreana apresenta uma situação semelhante ao arquipélago japonês, onde a língua principal da região é frequentemente vista como isolada. A correta identificação de uma família coreânica ocorre somente considerarmos as línguas de Yukjin e Jeju como de fato línguas independentes, e não como dialetos (uma visão tradicional encontrada em Yeon (2012) e Sohn (1999)) (CHO e WHITMAN, 2020, p. 13; LEE e RAMSEY, 2011, p. 3; GEORG, 2018, p. 151). Outro ponto de semelhança é o iminente risco de extinção dessas línguas, com a língua de Jeju limitada a cerca de 5 a 10 mil falantes, cuja fluência pode não ser suficiente para manter uma conversa (YANG, YANG e O'GRADY, 2020, p. 11). Quanto à língua de Yukjin, parece não haver levantamentos sobre o número de falantes, com os dados sendo omitidos em descrições recentes (KIM, 2003; KWAK, 2012; PIAO, 2019).

Em resumo, assim como no Japão, as línguas de Okinawa são frequentemente consideradas dialetos, mesmo sendo ininteligíveis para outros falantes da língua japonesa. Situação semelhante ocorre no continente.

As similaridades e diferenças entre as formações linguísticas do arquipélago japonês e da península coreana foram explicadas em detalhes em Tanaka de Lira (No Prelo), mas os detalhes estão omitidos aqui devido às limitações de espaço. Em resumo, os motivos explicados no artigo acima para as frequentes comparações entre as famílias coreânica e japônica são os seguintes:

**Desconhecimento das relações genéticas:** como já mencionado, embora haja várias hipóteses que conectam as famílias coreânica e japônica, especialmente em conjunto com as línguas do centro-oeste asiático do tronco dito altaico, ainda há muitas incertezas (GEORG, 2018, p. 151), e não sabemos com certeza quais são as relações dessas línguas com outras línguas do mundo – uma situação que não é incomum na região do nordeste asiático (COMRIE, 1981).

**Migrações entre continente e o arquipélago:** sabe-se que, ao menos duas das três grandes migrações formadoras do povo japonês (as migrações Yayoi e Kofun), se não foram originárias da península coreana, ao certo passaram por ela para chegar ao Japão. E, no sentido contrário, há a presença de genoma Jômon, tradicionalmente ligado ao Japão, em territórios da antiga Confederação de Gaia, cuja relação com o Japão é mencionada até mesmo no Nihon Shoki. Apesar de ser tratada no registro como uma possível colônia, e haver referência a uma invasão de Silla pela lendária imperatriz Jingû, há presença de DNA Jômon nesta região da península remetendo ao período (HENSHALL, 2004, p. 17; GELABERT, BLAZYTE, *et al.*, 2022).

**Possível presença de Línguas Japônicas no Continente:** Remetendo ao Período dos Três Reinos, do século IV ao século VI registrado do Samguk Sagi (SETH, 2011), há registros de possíveis cognatos japônicos no extrato lexical do reino coreano de Goguryeo, o mais distante do arquipélago japonês. Entre os possíveis candidatos estão as palavras “3” \*mil; “5” \*wucha; “7” \*nanun; “10” \*tek; “vale” \*tan, \*twon e \*thon; “coelho” \*wosaham e “chumbo” \*namwul (LEE e



RAMSEY, 2011, p. 41-43), e o texto inscrito na Estela de Gwanggaeto, atualmente na China, é uma das fontes preservadas da língua do reino. Obras como Beckwith (2004), apresentam um extenso vocabulário de possíveis palavras cognatas com o japonês no reino de Goguryeo.

**Convergências gramaticais:** como mencionado em Tanaka de Lira (2021), tanto o arquipélago japonês quanto a península coreana fazem parte de uma zona de convergência linguística, apresentando características tipológicas bastante similares, algumas das quais são bastante particulares à região, como a ênfase no uso de sujeito e tópico (LI e THOMPSON, 1976), enquanto outras são mais comuns, como classes de adjetivos que compartilham morfologia verbal (DIXON, 2010, p. 77-83; TANAKA DE LIRA, 2020). Ainda que essas semelhanças não indiquem uma origem comum, o histórico de contato tende a tornar as comparações entre as famílias linguísticas mais frequentes.

**Número de possíveis cognatos:** há séculos são feitas comparações entre as principais línguas das duas famílias, com Arai (1717) sendo um dos exemplos mais antigos conhecidos de análise preliminar de prováveis cognatos compartilhados pelas famílias coreânica e japônica. A primeira grande análise sistemática, presente em Martin (1966), sugeria 320 pares de cognatos, com diferentes graus de probabilidade, baseados em correspondências fonêmicas conforme a metodologia descrita na seção seguinte.

Devido à frequente suposição de uma língua proto-Nipo-Coreana, que seria ancestral das famílias coreânica e japônica (muitas vezes inserida em um tronco maior, como hipotetizado por Robbeets, Bouckaert, et al. (2021)), é necessário apresentar, em língua portuguesa, uma descrição mais detalhada dos argumentos que apoiam a existência dessa protolíngua, bem como os desafios relacionados à aceitação das reconstruções propostas. Para isso, será primeiro fornecido um resumo introdutório dos conceitos básicos em linguística histórica. Em seguida, serão listadas as mudanças fonéticas propostas para sugerir a existência do proto-Nipo-Coreano. Por fim, será realizada uma análise dos problemas associados a essas mudanças genéticas, seguida de uma conclusão com possíveis alternativas para os desafios apresentados.

## 2. Metodologia

A ferramenta mais amplamente aceita na linguística histórica, que tem como objetivo formular hipóteses de maneira científica e reconstruir o inventário segmental e suprasegmental, incluindo fonologia, morfologia, sintaxe e léxico, de uma língua ancestral e seus reflexos em suas descendentes, com base nos dados fornecidos por estas últimas, é o Método Comparativo (KIPARSKY, 2015, p. 65; WEISS, 2015, p. 127; HALE, 2015, p. 146; RANKIN, 2003). Isso significa que a metodologia utilizada aqui baseia-se na análise dos dados das línguas documentadas e, com base nesses dados, formula hipóteses e reconstrói como seria

a língua ancestral. A língua da qual as línguas descendentes surgiram é chamada de protolíngua (CROWLEY e BOWERN, 2010, p. 6; CAMPBELL, 1998, p. 108-9), e os cognatos encontrados nas línguas atuais, reflexos.

Isso implica que algumas precauções devem ser tomadas.

A primeira delas refere-se aos empréstimos linguísticos que são conhecidos por serem recentes nessas línguas. Apesar de serem tipologicamente mais próximas das línguas siberianas do Norte do que das línguas siníticas do Sul, a China tem sido, nos últimos milênios, uma fonte de conceitos culturais e palavras para a região, introduzindo práticas agrícolas e a escrita (HENSHALL, 2004, p. 10; SETH, 2011, p. 4). Por sua vez presença de povos cujas culturas deu origem às sociedades atuais na península e no arquipélago pré-datam esse contato em pelo menos três outros milênios no continente (SETH, 2011, p. 11), se tomarmos a cultura Jeulmun como ponto de partida, ou até mesmo outros dez milênios se tomarmos como ponto de partida a chegada do povo Jômon (HENSHALL, 2004, p. 9).

Isso significa que existem vários empréstimos de origem chinesa nas línguas desses povos, e a reconstrução desses reflexos é perfeitamente possível devido à recente adoção desses vocábulos, como demonstrado no quadro abaixo:<sup>5</sup>:

Significado	Mandarim	Wu	Coreano	Japonês	CN Med
País	/guó/	/koʔ/	/kuk/	/koku/	*kwək̚
Pessoa	/rén/	/nyin/	/in/	/nin/	*ŋin
Um	/yī/	/iʔ/	/il/	/iti/	*ʔit̚
Dois	–	/nyi/	/i/	/ni/	*ŋi
Três	/sān/	/se/	/sam/	/san/	*sam

Quadro 1. Comparação de palavras em línguas mandarim e wu e seus empréstimos em coreano e japonês, com uma reconstrução de chinês medieval para fins de demonstração. Os dados das línguas chinesas são retirados de Norman (2003), Zhengzhang e Zheng (2015) e Sun (2006). O fato de serem palavras independentes nas línguas de origem não implica independência fonológica ou semântica nas línguas que as adotaram.

Com base no quadro acima, é possível tirarmos algumas mudanças regulares que poderiam erroneamente ser tomadas como argumento para origem comum das línguas:

5 O quadro é oferecido com algumas simplificações por questões de espaço: a palavra dois em mandarim, /èr/, se trata de uma inovação e está sendo desconsiderada por não ser cognata com as outras formas presentes no quadro, e no caso da língua japonesa, quando a língua apresenta de forma produtiva tanto uma forma kan-on, de origem do mandarim, e uma origem go-on, de origem da língua Wu (MIYAKE, 2003, p. 104), apenas a última foi selecionada.

Significado	Mandarim	Wu	Coreano	Japonês
*k-	/g/	/k/	/k/	/k/
*-k	∅	/ʔ/	/k/	/ku/
*n-	/r/	/n/	∅	/n/
*-n	/n/	/n/	/n/	/N/
*s-	/s/	/s/	/s/	/s/
*-m	/n/	∅	/m/	/N/
*-t	∅	/ʔ/	/l/	/ti/

Quadro 2. Regras de mudança fonética entre as línguas mandarim, wu, coreana e japonesa com base no Quadro 1.

Ainda que o quadro acima seja insuficiente mesmo se de fatos as quatro línguas tivessem a mesma origem, ele é suficiente para mostrar algumas regularidades. Focando apenas nas consoantes, como a língua japonesa não permite oclusivas (/k/ e /t/), houve a inserção de uma vogal epentética, assim como a distinção entre /n/ e /m/ finais foi perdida, havendo apenas uma nasal em posição de coda /N/. No caso da língua coreana, a diferença foi que houve o apagamento de \*n no início de sílaba, levando às formas atuais de pessoa (/in/) e dois (/i/). Além de, é claro, ser possível constatar que as línguas chinesas não continuaram inertes, com o número 2 sendo sequer cognato em mandarim e wu.

Se, de forma mais apropriada para análise de origem genética, utilizarmos não os empréstimos, mas as palavras nativas originais das línguas japonesa e coreana para os termos do Quadro 1, a improbabilidade se haver uma origem comum se torna mais aparente:

Significado	Mandarim	Wu	Coreano Antigo	Japonês Antigo
País	/guó/	/koʔ/	/nalak/	/kuni/
Pessoa	/rén/	/nyin/	/salɒm/	/pitə/
Um	/yī/	/iʔ/	/hɒnah/	
Dois	/əɾ/	/nyi/	/dulh/	/puta/
Três	/sān/	/se/	/seyh/	/mit/

Quadro 3. Reformulação do com as palavras originais do coreano e do japonês antigo, retirados de Frellesvig (2010), Lee e Ramsey (2011) e Vovin (2020).

*A prima facie*, não há uma correspondência sequer entre as línguas. É possível supor que /salɒm/ e /pitə/ sejam cognatos se supormos algumas mudanças fonêmicas, por exemplo, mas essas mudanças não podem ser feitas de maneira *ad hoc*. Por exemplo, se (1) supormos que /pitə/ originalmente se referiria a pessoa em proto-nipo-coreano, é perfeitamente possível que formulemos a hipótese de que \*-t- se manteve /-t-/ em

japonês antigo, mas se tornou /-l-/ em coreano: isso explicaria o porquê de termos /t/ em japonês e /l/ em coreano nas palavras “pessoa” e “dois”. Entretanto, precisaríamos explicar o porquê de as palavras possuírem a mesma consoante inicial nas línguas do arquipélago (/p-/) mas não na península (/s-/ e /d-/ respectivamente).

Isso não quer dizer que o cuidado com empréstimos se limite às línguas de fora da região. Obras como Vovin (2010) e Whitman (2012), que visam, dentre outros objetivos, sugerir correspondências fonêmicas para reconstrução da língua proto-nipo-coreana, têm como mérito o fato de corrigirem o que foi percebido posteriormente como sendo algumas falhas metodológicas de Martin (1966).

Um exemplo de refinamento metodológico foi considerar que, se um cognato está disponível apenas no Japonês Antigo Ocidental, ou seja, nos dialetos mais próximos da península coreana naquela época, então as chances de esse termo ser um empréstimo resultante de uma inovação na língua coreana são maiores do que a probabilidade de ser um vocábulo de origem comum às duas famílias. Além disso, muitas vezes foram usadas formas contemporâneas das línguas, quando o ideal seria terem sido utilizadas as formas mais antigas das quais temos registros.

Feitas essas ressalvas em relação à metodologia, é possível entender as dificuldades com os dados das línguas: a falta de regras de mudança fonêmica consistentes para reconstrução de um proto-nipo-coreano.

### 3. Dados e Análise

Tendo em mente as devidas precauções mencionadas na seção anterior, o primeiro passo ao coletar os dados foi procurar por um conjunto sistemático de cognatos que pudessem ser comparados como parte de um sistema. Certos termos tendem a ser mais resistentes a empréstimos, como numerais baixos, partes do corpo e termos geográficos (CAMPBELL, 1998, p. 112). Portanto, a primeira decisão foi buscar por cognatos entre numerais e pronomes interrogativos: ambos os grupos podem ser vistos como parte de um sistema e são resistentes a empréstimos – mesmo no caso das línguas da região, em que houve um empréstimo de números sinóticos, os sistemas numéricos nativos continuam a ser usados simultaneamente.

No entanto, antes de apresentar os dados, é necessário fazer algumas observações sobre a escrita e a fonologia reconstruída do japonês clássico e do proto-japonês devido à sua importância na família japônica. Seguindo as reconstruções do japonês antigo e do proto-japonês por Bentley (2012), Frellesvig (2010), Miyake (2003) e Vovin (2020), é possível estabelecer com certo grau de confiança que, ao contrário da língua contemporânea, a língua antiga possuía cerca de seis a oito fonemas vocálicos, com alguma variação na identificação de quais seriam monotongos e quais seriam ditongos. Essas discrepâncias surgem porque as reconstruções são baseadas no uso de diferentes ideogramas em documentos como o *Man'yōshū* e o *Kojiki*, e na comparação com as leituras desses ideogramas em japonês e em outras línguas contemporâneas, como as

línguas chinesas, coreano e vietnamita. O modelo mais comum entre as obras citadas, com sete monotongos e um ditongo, apresenta um quadro semelhante ao exposto abaixo:

i /i/	i <sub>2</sub> /i̥/	u /u/
e <sub>1</sub> /e/	e <sub>2</sub> /əj/	o <sub>1</sub> /ə/
	a /a/	o <sub>2</sub> /o/

Quadro 4. Inventário de fonemas vocálicos da língua japonesa antiga, com grafemas tradicionais e valor fonêmico (aqui utilizado).

Em relação às consoantes, o inventário apresentado pelas obras é mais consistente, com o seguinte quadro sendo amplamente aceito para reconstrução do japonês antigo:

	Bilabial	Alveolar/Dental	Palatal	Velar
Obstr. Surdos	/p/ [p]	/t/ [t]    /s/ [s]		/k/ [k]
Pré-Nasalizadas	/b/ [ᵐb]	/d/ [ᵐd]    /z/ [ᵐz]		/g/ [ᵑg]
Nasais	/m/ [m]	/n/ [n]		
Glide e Vibrante	/w/ [w]	/r/ [ɹ]	/y/ [y]	

Quadro 5. Inventário de fonemas consonantais da língua japonesa antiga, com valor fonêmico e valores fonéticos reconstruídos.

A maior divergência se refere à produção fonética das consoantes surdas, com Frellesvig (2010, p. 35) propondo espirantização das consoantes (exceto /t/) em variação livre, além de sonorização em meio de palavra, com /p/, para citar um exemplo, podendo ser realizado como [p~ϕ] em início de palavra e [b~β] em meio de palavra. Como esse fator não parece ter impacto na reconstrução de etapas anteriores da língua, não será considerado relevante aqui.

A situação para reconstrução do coreano antigo não é muito diferente, com os mesmos métodos sendo utilizados para se supor qual seria o inventário da língua. Obras como Lee e Ramsey (2011) e Nam (2012) reconstróem um sistema com sete vogais, havendo um consenso neste particular. No caso da vogal /ʌ/, não há confirmação se a vogal era labializada ou não, ou se ela era centralizada motivo pelo qual por vezes também é representada como /ɔ/ (LEE e RAMSEY, 2011, p. 67) ou /ɐ/ (NAM, 2012, p. 64), mas independentemente de seu valor fonético, o inventário fonêmico se assemelha ao seguinte:

/i/ [i]	/i/ [i]	/u/ [u]
	/ə/ [ə]	/o/ [o]
	/a/ [a]	/ʌ/ [ɐ~ʌ~ɔ]

Quadro 6. Inventário de fonemas vocálicos da língua coreana antiga, com valor fonêmico e valores fonéticos reconstruídos.

As consoantes apresentam um quadro um pouco mais claro, excetuando-se pela omissão de possíveis glides na língua que, por não haver pesquisas conclusivas confirmando sua existência, não podem por ora serem incluídas no inventário consonantal (LEE e RAMSEY, 2011, p. 68). Mesmo desconsiderando essa lacuna, é possível reconstruir um quadro como o abaixo:

	Bilabial	Alveolar/Dental	Palatal	Velar / Glotal
Ocl.	/b/ [p]	/d/ [t]	/j/ [ɟ]	/g/ [k]
Ocl. Aspiradas	/p/ [p <sup>h</sup> ]	/t/ [t <sup>h</sup> ]	/c/ [tʃ <sup>h</sup> ]	/k/ [k <sup>h</sup> ]
Nasais	/m/ [m]	/n/ [n]		/ŋ/ [ŋ]
Fricativas		/s/ [s]	/z/ [z]	/h/ [h]
Líquidas		/l/ [l]	/r/ [r]	

Quadro 7. Inventário de morfemas consonantais da língua coreana antiga, com valor fonêmico e valores fonéticos reconstruídos.

*A prima facie*, a maior diferença na fonologia de ambas as línguas seria a presença de consoantes aspiradas em coreano, o que apontaria para uma incompatibilidade comparativa com as línguas do arquipélago japonês, e a existência de um contraste de sonoridade no japonês, inexistente na península coreana.

Entretanto, assim como no caso da língua japonesa se sabe que as consoantes sonorizadas surgiram de processos de pré-nasalização, também é sabido que isso se deveu a casos em que ou houve síncope de vogais entre /h/ e a consoante seguinte, ou a aspiração se encontrava em posição de coda na sílaba e passou por um processo de metátese. Como exemplo de cada processo, a palavra para grande, [k<sup>h</sup>i-], veio de originalmente de [hiki], enquanto [koh], nariz, resultou em [k<sup>h</sup>o] (LEE e RAMSEY, 2011, p. 64-65). Pelo lado da língua japonesa, a presença do contraste entre consoantes surdas e sonoras também é conhecidamente o resultado de assimilação de um processo de pré-nasalização, e que ainda é claro nas mudanças de paradigma verbal do japonês antigo para o japonês moderno, como na queda de contraste entre [yomite] (1er.CONJ) e [ywo<sup>mb</sup>bite] (chamar.CONJ) para [yonde] (FRELLESVIG, 2010, p. 35).

Além disso, com alguma frequência é (corretamente) levado em consideração o acento tonal nas duas línguas ao se fazer reconstruções e identificações de empréstimos, dado o longo contato. Para fins de simplicidade, essa variável será ignorada nas próximas páginas.

Uma vez explicadas as características fonológicas das antigas línguas japonesa e coreana, bem como os valores representados por cada um dos símbolos utilizados a seguir, é possível avançar para a comparação lexical, buscando por cognatos e correspondências entre famílias.

Para melhor compreensão dos processos evolutivos dentro da família japônica e da língua coreânica, foram consultados dados antigos de outras línguas da família. Como nem sempre foram encontradas fontes sobre versões antigas das línguas Ryûkyû, os dados da língua okinawana foram obtidos a partir de Loveless (1963, p. 102), sem referência às outras línguas do ramo. Para a língua japonesa, foram utilizados os dados disponíveis Vovin (2020) e Frellesvig (2010), originalmente retirados de fontes escritas em Man'yôgana (UNIVERSITY OF VIRGINIA LIBRARY ELECTRONIC TEXT CENTER, 759; Ô, 2014). Os dados da língua coreana, da mesma forma, são originários de Kim (1145) e Samguk Yusa e cruzados com Lee e Ramsey (2011) e Vovin (2010), com referência exclusiva às últimas obras quando não foi possível obtenção de outra forma. No caso da família coreânica, nem sempre foram obtidos dados confiáveis das outras línguas da família, não tendo, portanto, sido referenciados nas comparações abaixo.

Em resumo, temos o seguinte quadro ao se comparar as famílias coreânica e japônica:

Família Coreânica	Família Japônica		
Coreano Clássico	Japonês Clássico	Okinawano Atual	Tradução
/misik/	/nani/	/nu:/	Que
/nu/	/tare/	/ta:/	Quem
/əntuj/	/idu(ku)/	/ma:/	Onde
/əntʃej/	/itu/	/itʃi/	Quando
/ənu/	/indu(re)/	/jiru/	Qual?

Quadro 8. Pronomes interrogativos em coreano e japonês clássico e em okinawano contemporâneo.

O primeiro cuidado que deve ser observado ao analisar o quadro acima é reconhecer que existem algumas inovações no okinawano, como a palavra para “onde”, que se acredita ter sua origem em \*idu- “onde” e \*-ma “lugar”, ainda sendo usada nas duas partes remanescentes no dialeto de Tonoshiro da ilha de Ishigaki (VOVIN, 2020, p. 306). No caso das palavras “quando” e “quem”, não há motivo para acreditar que as palavras não sejam cognatas.



Um possível paralelo que pode sugerir uma origem comum entre os pronomes interrogativos das duas línguas está relacionado à constância com que outras palavras interrogativas em coreano começam com \*ən-, enquanto as mesmas palavras em japonês apresentam \*id- e \*it-. Levando em consideração que a vocalização de consoantes em japonês ocorreu após um fenômeno de pré-nasalização em sua história e supondo que nesses ambientes de fato havia uma consoante nasal no proto-japonês, a hipótese de que \*ən- em proto-coreano e \*in- em proto-japonês seriam reflexos de uma mesma palavra interrogativa caso a hipótese do proto-nipo-coreano seja verdadeira. No entanto, geralmente a nasal não é considerada parte da morfologia original (THORPE, 1983; VOVIN, 2020; WHITMAN, 2012), o que dificulta essa hipótese. Também existe a possibilidade de que essa seja uma característica areal, visto que pronomes interrogativos de línguas vizinhas, como a língua tungúsica manchú, frequentemente começam com e- ou ai- (GORELOVA, 2002, p. 219-220), mas tal hipótese ainda precisaria de melhor embasamento.

Mesmo se excluirmos o papel do contato e nos concentramos apenas na possibilidade de uma origem comum, enfrentariamos duas dificuldades. A primeira seria estender essa regra para outros contextos e vocábulos. Como demonstrado anteriormente, o método comparativo não restringe a busca por mudanças fonêmicas apenas a uma classe lexical. Seguindo essa linha de raciocínio, Whitman (2012, p. 30-31) apresenta alguns pares de reflexos nas línguas proto-coreana e proto-japonesa, como \*jə e \*i, respectivamente, para palavras como \*hjə- e \*sir- (“branco”) e \*mjəʃ- e \*mit- para “três/alguns”. No entanto, seria necessário explicar a presença do glide inicial nesses contextos e sua ausência nas palavras interrogativas. Além disso, se esse for o caso, precisaríamos explicar por que a nasalização desapareceu na palavra “quando” em proto-japonês, caso seja realmente uma palavra bimorfêmica, como defendido por Thorpe (1983, p. 223). Caso contrário, apenas teríamos \*ə- e \*i- como morfema inicial nessas línguas para as outras palavras, seguindo Vovin (2020, p. 304, 310).

Também poderia ser verossímil, através de processos de deriva semântica, que \*misik em coreano clássico e \*mosi em japonês tivessem uma origem comum, como proposto por Whitman. No entanto, como já apontado por Vovin (2010, p. 199), essa palavra não é encontrada em textos produzidos mais a leste do arquipélago japonês (o que poderia indicar um possível empréstimo) e também nunca é usada para formular perguntas. Dada a persistência ao longo de diferentes famílias linguísticas dos pronomes interrogativos (como é o caso de “que” em português, “ke” em nepali e “chto” em russo, que correspondem às regras de correspondência fonêmica entre essas línguas e o proto-indo-europeu), essa hipótese ainda carece de argumentos mais sólidos.

O quadro com os numerais apresenta uma situação ainda mais complexa:

Família Coreânica	Família Japônica		
Coreano Clássico	Japonês Clássico	Okinawano Atual	Tradução
/hʌnah/	/pitə/	/ti:/	1
/dulh/	/puta/	/ta:/	2
/seh/	/mi/	/mi:/	3
/neh/	/yə/	/yu:/	4
/dasʌs/	/itu/	/ʔitʃi/	5
/jəsis/	/mu/	/mu:/	6
/nilgob/	/nana/	/nana/	7
/jədilb/	/ya/	/ya:/	8
/ahob/	/kəkənə/	/kukunu/	9
/jəlh/	/təwo/	/tu:/	10

Quadro 9. Numerais de 1 a 10 em coreano e japonês clássicos e em okinawano contemporâneo.

Em um primeiro olhar, é perfeitamente possível identificar as correspondências entre o japonês clássico e o okinawano. Com exceção dos dois primeiros numerais, onde é evidente que ocorreu um processo de haplologia que reduziu o número de sílabas (ainda que não de moras), é viável traçar correspondências entre as consoantes: JP: OK, m:m, t:t, y:y, n:n, k:k, Ø:ʔ e w\_ø:Ø\_u.

No que diz respeito às vogais, \*a em japonês clássico sempre corresponde a /a/ em oki-nawano. Houve uma elevação das vogais médias \*ə e \*o do japonês clássico, que atualmente correspondem ao fonema /o/ no japonês contemporâneo, em okinawano, resultando em /kukunu/ e /tu:/, em comparação a /kokono/ e /to:/ no japonês contemporâneo. Também é notável algumas inovações no okinawano, como o alongamento vocálico para tornar todos os numerais, ao menos, bimoráicos. Por isso, os dois primeiros numerais no okinawano, mesmo após o processo de haplologia, possuem o mesmo número de moras. Além disso, \*tu, em posição final de palavra, regularmente resultou em /ʃi/ em okinawano, como se pode observar não apenas no numeral 5, mas também na palavra “quando”, disponível para referência no Quadro 8.

Enquanto é possível reconhecer os cognatos dentro da família japônica, não há aparente correspondência entre os vocábulos do coreano clássico e do japonês clássico em nenhum dos numerais. Essa falta aparente de correspondência não é algo novo, uma vez que essa dificuldade, especialmente quando combinada com comparações com as línguas consideradas altaicas, tem sido amplamente discutida por Ramer e Sidwell (1997), os quais consideram a ausência de cognatos como irrelevante. Ainda assim, há várias explicações para as diferenças entre os defensores de uma origem comum entre

as duas línguas e/ou as línguas consideradas altaicas, desde as abordagens de Miller (1969), Hamp (1970) e Blažek (1999).

No caso da língua japonesa, há muito tempo se sabe que os números são dispostos em pares (1 e 2, 3 e 6, 4 e 8), com mudanças regulares nas vogais, como de /o/ para /a/ (ELLIS, 1873, p. 50). Tem sido comum as comparações partirem dos seguintes pressupostos: que houve um processo de inovação que levou à divergência das línguas, comparando \*pitə “um” do proto-japonês com \*piris “primeiramente, começar” em coreano medieval e \*puta “dois” do proto-japonês com \*pʃak “par” em coreano medieval (WHITMAN, 2012, p. 33; BLAŽEK, 1999, p. 57), apesar de Miller (1969) estabelecer cognatos divergentes usando o ablaut já mencionado. Acima do número quatro, simplesmente não há suposições conhecidas de pares entre as famílias coreânica e japônica até o momento. A comparação de \*p para proto-japonês e proto-coreano, baseada apenas nesses dois pares, não seria convincente para estabelecer uma relação entre as duas línguas (e, de fato, nenhum dos autores das obras anteriores limitou-se a esses exemplos, sendo que os outros vocábulos utilizados estão fora do escopo da comparação aqui empregada). Embora o número 7 comece com o mesmo fonema tanto no coreano quanto no japonês, não há razão aparente para considerá-lo um par de cognatos, e é difícil elaborar regras de correspondência que justifiquem a associação entre /nilgob/ e /nana/ como reflexos de um ancestral comum.

É possível que as línguas tenham uma origem comum sem que haja cognatos em seus numerais e pronomes, embora, como mencionado anteriormente, essas palavras sejam mais resistentes a empréstimos (CAMPBELL, 1998, p. 112). Com base somente nessas palavras, no entanto, teríamos as seguintes correspondências:

Proto-Coreano	Proto-Japonês	Proto-Nipo-Coreano
*ə-	*i-	*e-
*p-	*p-	*p-

Quadro 10. Reconstruções de correspondências fonêmicas em proto-nipo-coreano com base apenas nos numerais cardinais e pronomes interrogativos das línguas das famílias coreânica e japônica

Como mencionado anteriormente, as correspondências acima, deduzidas dos dados exibidos anteriormente, apresentam suas complexidades. No primeiro caso, ao considerarmos um alçamento da vogal em japonês e sua centralização em coreano, surge a necessidade de compreender o comportamento da nasal seguinte, que está presente em todos os dados, exceto \*itu (quando) em proto-japonês. A correspondência entre as línguas de uma vogal de origem comum nos pronomes interrogativos poderia ser um argumento favorável à existência de um ancestral comum entre as duas línguas.

Se aceitarmos o argumento anterior de que pode haver parentesco linguístico com base na possível presença de cognatos no sistema pronominal, a falta aparente de cognatos nos numerais, no entanto, embora não seja uma evidência conclusiva de uma não-relação entre as línguas, poderia, no mínimo, levar a pelo menos uma das duas possibilidades seguintes:

- (a) A separação entre os falantes ocorreu em um período distante o suficiente para que, desde então, ocorressem inovações no sistema numérico de pelo menos uma das famílias;
- (b) Houve uma influência do substrato no arquipélago japonês ou na península coreana, que pode ter sido a fonte dos numerais, mas não dos pronomes;

Argumentos podem ser apresentados para ambas as possibilidades.

Na possibilidade (a), a separação, conforme cálculos de inferência bayesiana presentes em Robbeets (2020, p. 38), ocorreu em torno de 2000 a 3000 AC, o que significa que não temos registros da história das famílias por um período de aproximadamente três a quatro mil anos, se tomarmos como ponto de partida os primeiros documentos ainda existentes das línguas mais faladas. Apesar de existirem ferramentas para o cálculo da reposição de palavras, como a glotocronologia, muitas das premissas dessa abordagem já foram rejeitadas ou contestadas pela linguística contemporânea (CAMPBELL, 1998, p. 314), não havendo meios exatos para definir a probabilidade de uma mudança completa do sistema numérico dessas línguas no período disponível. Por esse motivo, como mencionado anteriormente, foi proposta a hipótese de uma série de inovações no sistema numérico dessas famílias (BLAŽEK, 1999; HAMP, 1970; MILLER, 1969).

De qualquer forma, considerando a proximidade geográfica e o contato frequente entre os falantes das duas línguas, caso haja uma origem comum, seria necessário explicar como essas inovações ocorreram de forma independente, ou seja, sem evidências claras de empréstimo linguístico de uma língua para outra, e por que tem sido tão desafiador estabelecer pares confiáveis de cognatos ao longo dos últimos quase 80 anos (VOVIN, 2010, p. 3).

Na possibilidade (b), o histórico de diferentes migrações no arquipélago japonês, como a Jômon, a Yayoi e a Kofun (HENSHALL, 2004; COOKE, MATTIANGELI, *et al.*, 2021), e a possibilidade dos três reinos formadores da península coreana falarem diferentes línguas ou serem multiétnicos com vastas partes da população falando dialetos de uma mesma língua (BECKWITH, 2005; JANHUNEN, 2005; UNGER, 2005), poderiam explicar a divergência nos numerais nas línguas. A presença de numerais cognatos aos japoneses na península coreana, como mencionado a respeito da Estela de Gwanggaeto, sugere que em um período não muito distante à chegada da escrita na península coreana, ainda havia o uso de um sistema numérico similar ao presente hoje na família japônica.

## Conclusão

Devido às restrições de espaço e à complexidade da atividade proposta, não é possível aprofundar ainda mais nos desafios relacionados à busca de correspondências fonêmicas entre as famílias coreânica e japônica para estabelecer a existência de uma protolíngua comum às duas famílias, aqui chamada tentativamente de proto-nipo-coreano.

Essa discussão serve como uma amostra das dificuldades envolvidas na identificação de cognatos e no estabelecimento de regras de correspondência que permitam reconhecer como possíveis reflexos em ambas as línguas teriam surgido de um ancestral comum. Esse desafio em si torna as hipóteses sobre a existência de troncos linguísticos mais abrangentes, que englobariam ambas as famílias, ainda mais incertas.

Apesar das dificuldades em identificar cognatos para estabelecer correspondências entre as famílias linguísticas, é necessário buscar uma explicação para as semelhanças quando essas não podem ser atribuídas apenas ao contato linguístico.

A busca pelas origens das línguas na região é de extrema importância, não apenas para a linguística histórica, a fim de tentar preencher as lacunas mencionadas ao longo do texto, mas também para um melhor entendimento da formação do nordeste asiático e das relações e histórias dos povos coreano e japonês nesse contexto global que chamamos de lar.

## Bibliografia

- ARAI, H. Tōga. **Internet Archive**, 1717. Disponível em: <<https://archive.org/details/JAPAF.1219.1/page/n7/mode/2up>>. Acesso em: 20 jun. 2022.
- BARBER, A.; RAMIREZ, E. G. Idiolects. **The Stanford Encyclopedia of Philosophy (Summer 2021 Edition)**, 2021. Disponível em: <<https://plato.stanford.edu/archives/sum2021/entries/idiolects/>>. Acesso em: 28 out. 2022.
- BECKWITH, C. I. **Koguryo: The Language of Japan's Continental Relatives**. Leiden, NL: Brill, 2004.
- BECKWITH, C. I. The Ethnolinguistic History of the Early Korean Peninsula Region: Japanese-Koguryōic and other Languages in the Koguryō, Paekche, and Silla kingdoms. **Journal of Inner and East Asian Studies**, Seoul, v. 2, n. 2, p. 33-64, December 2005.
- BENTLEY, J. R. Old Japanese. In: TRANTER, N. **The Languages of Japan and Korea (Routledge Language Family Series)**. Oxon, UK: Routledge, 2012. p. 189-211.
- BLAŽEK, V. Altaic Numerals. In: BLAŽEK, V. **Numerals: Comparative-Etymological Analyses of Numeral Systems and Their Implications (Saharan, Nubian, Egyptian, Berber, Kartvelian, Uralic, Altaic and Indo-European Languages)**. Brno: Masarykova Univerzita, 1999. p. 102-140.
- BOWERN, C.; EVANS, B. **The Routledge Handbook of Historical Linguistics**. New York, NY: Routledge, 2015.
- BUGAEVA, A. Southern Hokkaido Ainu. In: TRANTER, N. **The Languages of Japan and Korea (Routledge Language Family Series)**. Oxon, UK: Routledge, 2012. p. 461-509.

- CAMPBELL, L. **Historical Linguistics: An Introduction**. Cambridge, MA: The MIT Press, 1998.
- CAMPBELL, L. Language Isolates and Their History. In: CAMPBELL, L. **Language Isolates (Routledge Language Family Series)**. Oxon, UK: Routledge, 2018. p. 1-18.
- CHO, S.; WHITMAN, J. **Korean: A Linguistic Introduction**. New York, NY: Cambridge University Press, 2020.
- CLACKSON, J. **Indo-European Linguistics**. New York, NY: Cambridge University Press, 2007.
- COATES, J. Gender. In: LLAMAS, C.; MULLANY, L.; STOCKWELL, P. **The Routledge Companion to Sociolinguistics**. New York, NY: Routledge, 2007. p. 62-68.
- COMRIE, B. **The Languages of the Soviet Union**. Cambridge, UK: Cambridge University Press, 1981.
- COOKE, N. P. et al. Ancient Genomics Reveals Tripartite Origins of Japanese Populations. **Science Advances**, Washington, DC, v. 7, n. 38, p. 1-15, September 2021.
- COUNCIL OF LOCAL AUTHORITIES FOR INTERNATIONAL RELATIONS. **Local Government in Japan**. Tokyo, JP: Council of Local Authorities for International Relations (CLAIR), 2004.
- CROWLEY, T.; BOWERN, C. **Historical Linguistics**. New York, NY: Oxford University Press, 2010.
- DE BOER, E. The Classification of the Japonic Languages. In: ROBBEETS, M.; SAVELYEV, A. **The Oxford Guide to the Transeurasian Languages**. Oxford, UK: Oxford University Press, 2020. p. 40-58.
- DIXON, R. M. W. **Basic Linguistic Theory: Volume 2 - Grammatical Topics**. Oxford, UK: Oxford University Press, 2010.
- DOUGHERTY, T. Ainu. In: CAMPBELL, L. **Language Isolates (Routledge Language Family Series)**. Oxon, UK: Routledge, 2018. p. 100-116.
- EBERHARD, D. M.; SIMONS, G. F.; FENNIG, C. D. What are the largest language families? **Ethnologue: Languages of the World**, 2022. Disponível em: <<https://www.ethnologue.com/guides/largest-families>>. Acesso em: 25 out. 2022.
- ELLIS, R. **On Numerals: As Signs of Primeval Unity Among Mankind**. London, UK: Trübner & Co., 1873.
- FRANÇOIS, A. Trees, Waves and Linkages: Models of Language Diversification. In: BOWERN, C.; EVANS, B. **The Routledge Handbook of Historical Linguistics**. New York, NY: Routledge, 2015. p. 161-189.
- FRELLESVIG, B. **A History of the Japanese Language**. Cambridge, UK: [s.n.], 2010.
- GELABERT, P. et al. Northeastern Asian and Jomon-related genetic structure in the Three Kingdoms period of Gimhae, Korea. **Current Biology**, Maryland Heights, MO, v. 32, p. 1-13, August 2022.

- GEORG, S. Other Isolated Languages of Asia. In: CAMPBELL, L. **Language Isolates (Routledge Language Family Series)**. Oxon, UK: Routledge, 2018. p. 139-161.
- GEORG, S. Other Isolated Languages of Asia. In: CAMPBELL, L. **Language Isolates (Routledge Language Family Series)**. Oxon, UK: Routledge, 2018. p. 139-161.
- GORELOVA, L. M. **Manchu Grammar**. Leiden, NL: Koninklijke Brill, 2002.
- GRIMES, B. F. **Ethnologue: Languages of the World**. 12th. ed. Dallas, TX: Summer Institute of Linguistics, 1992.
- HALE, M. The Comparative Method: Theoretical Issues. In: BOWERN, C.; EVANS, B. **The Routledge Handbook of Historical Linguistics**. New York, NY: [s.n.], 2015. p. 146-160.
- HAMP, E. P. On The Altaic Numerals. In: JAKOBSON, R.; KAWAMOTO, S. **Studies in General and Oriental Linguistics Presented to Shirô Hattori on the Occasion of His Sixtieth Birthday**. Tokyo, JP: TEC, 1970. p. 188-197.
- HASEGAWA, Y. **Japanese: A Linguistic Introduction**. Cambridge, UK: Cambridge University Press, 2015.
- HEINRICH, P.; MIYARA, S.; SHIMOJI, M. Introduction: Ryukyuan Languages and Ryukyuan Linguistics. In: HEINRICH, P.; MIYARA, S.; SHIMOJI, M. **Handbook of the Ryukyuan Languages: History Structure and Use (Handbooks of Japanese Language and Linguistics)**. Berlin, DE: Walter de Gruyter, v. 11, 2015. p. 1-12.
- HENSHALL, K. G. **A History of Japan: From Stone Age to Superpower**. 2<sup>a</sup>. ed. New York, NY: Palgrave MacMillan, 2004.
- IWASAKI, S. **Japanese: Revised Edition**. Amsterdam, NL: John Benjamins Publishing Company, 2006.
- JANHUNEN, J. The Lost Languages of Koguryō. **Journal of Inner and East Asian Studies**, Seoul, v. 2, n. 2, p. 65-86, December 2005.
- JANHUNEN, J. A. Ainu Ethnic Origins. In: BUGAEVA, A. **Handbook of the Ainu Language**. Berlin, DE: Walter de Gruyter Inc, 2022. p. 57-78.
- JAPANESE GAZETEER. **Administrative Subdivisions of Japan, with Appendix of 47 Prefectural Maps**. Washington, DC: US Department of State. Division of Research for Far East., 1946.
- KAPOVIĆ, M. Indo-European Languages - Introduction. In: KAPOVIĆ, M. **The Indo-European Languages**. 2<sup>a</sup>. ed. New York, NY: Routledge, 2017. p. 1-12.
- KERSWILL, P. Social Class. In: LLAMAS, C.; MULLANY, L.; STOCKWELL, P. **The Routledge Companion to Sociolinguistics**. New York, NY: Routledge, 2007. p. 51-61.
- KIM, B. Samguk Sagi. 三國史記, 1145. Disponível em: <<https://db.history.go.kr/item/level.do?itemId=sg&types=o>>. Acesso em: 15 nov. 2022.
- KIM, S. H. 육진 방언의 종결어미 연구 (Yukjin Bagn'on ui Jonggyeoleomi Yeongu). **어문논집 (Eomun Nunjip)**, Iksan, Jeonbuk, KR, v. 48, p. 93-125, out. 2003.



- KIPARSKY, P. New Perspectives in Historical Linguistics. In: BOWERN, C.; EVANS, B. **The Routledge Handbook of Historical Linguistics**. New York, NY: [s.n.], 2015. p. 64-102.
- KWAK, C. 육진방언의 음성과 음운사 (Yukjin Bang'eon ui Eumseonggwa Eumunseo). **방언학 (Bang'eohag)**, Incheon, KR, v. 16, p. 121-154, dez. 2012.
- LAWRENCE, W. P. Southern Ryukyuan. In: TRANTER, N. **The Languages of Japan and Korea (Routledge Language Family Series)**. Oxon, OX: Routledge, 2012. p. 381-411.
- LEE, K.-M.; RAMSEY, S. R. **A History of the Korean Language**. Cambridge, UK: Cambridge University Press, 2011.
- LEE, S.; HASEGAWA, T. Bayesian Phylogenetic Analysis Supports an Agricultural Origin of Japonic Languages. **Proceedings of the Royal Society B**, London, UK, v. 278, p. 3662-3669, May 2011.
- LI, C. N.; THOMPSON, S. A. Subject and Topic: A New Typology of Language. In: LI, C. N. **Subject and Topi**. New York, NY: Academic Press, 1976. p. 457-489.
- LLAMAS, C. Age. In: LLAMAS, C.; MULLANY, L.; STOCKWELL, P. **The Routledge Companion to Sociolinguistics**. New York, NY: Routledge, 2007. p. 69-76.
- LOVELESS, O. **The Okinawan Language**. Ann Arbor, MI: University of Michigan, 1963.
- MARTIN, S. E. Lexical Evidence Relating Korean to Japanese. **Language**, Washington, DC, v. 42, n. 2, p. 185-251, 04-06 1966.
- MATSUMORI, A.; ONISHI, T. Japanese Dialects: Focusing on Tsuruoka and Ei. In: TRANTER, N. **The Languages of Japan and Korea (Routledge Language Family Series)**. New York, NY: Routledge, 2012. p. 313-348.
- MILLER, R. A. The Altaic Numerals and Japanese. **The Journal-Newsletter of the Association of Teachers of Japanese**, Pittsburgh, PA, v. 6, n. 2, p. 14-29, Outubro 1969.
- MIYAKE, M. Historical Sources and Periodization of the Japonic and Koreanic Languages. In: ROBBEETS, M.; SAVELYEV, A. **The Oxford Guide to the Transeurasian Languages**. Oxford, UK: Oxford University Press, 2020. p. 9-21.
- MIYAKE, M. H. **Old Japanese: A Phonetic Reconstruction**. London, UK: RoutledgeCurzon, 2003.
- MORAVCSIK, E. A. **Introducing Language Typology**. Cambridge, UK: Cambridge University Press, 2013.
- NAM, P.-H. Old Korean. In: TRANTER, N. **The Languages of Japan and Korea (Routledge Language Family Series)**. Oxon, UK: Routledge, 2012. p. 41-72.
- NORMAN, J. The Chinese Dialects: Phonology. In: THURGOOD, G.; LAPOLLA, R. J. **The Sino-Tibetan Languages**. London, UK: Routledge, 2003. p. 72-83.
- Ō, Y. N. **Kojiki (古事記・고사기)**. Seul, KR: Jimanji (Jishikeulmandeuneunjishik) Sasangseonjip, 2014.
- PIAO, M. 咸北 六鎮方言의 母音調和 實現 樣相에 대하여 - 세대별 差異를 中心으로 - (Hambuk Yukjin Bang'eon ui Moeum Johwa Silhyeon Yangsang e Daehayeo - Sedaebyeol

- Chai leul Jungsim euro -). **어문연구 (Eomun Yeongu)**, Seoul, KR, v. 47, n. 4, p. 67-87, dez. 2019.
- RAMER, A. M.; SIDWELL, P. The Altaic Debate and the Question of Cognate Numerals. **Wiener Zeitschrift für die Kunde des Morgenlandes**, Vienna, v. 87, p. 153-175, 1997.
- RANKIN, R. L. The Comparative Method. In: JOSEPH, B. D.; JANDA, R. D. **The Handbook of Historical Linguistics**. Malden, MA: Blackwell Publishing, 2003. p. 183-212.
- ROBBEETS, M. If Japanese is Altaic, How Can It Be So Simple? In: LUBOTSKY, A.; SCHAEKEN, J.; WIEDENHOF, J. **Evidence and Counter-Evidence: Essays in Honour of Frederik Kortlandt**. Amsterdam, NL: Rodopi, v. 2, 2008. p. 337-367.
- ROBBEETS, M. The Classification of the Transeurasian Languages. In: ROBBEETS, M.; SAVELYEV, A. **The Oxford Guide to the Transeurasian Languages**. Oxford, UK: Oxford University Press, 2020. p. 31-39.
- ROBBEETS, M. et al. Triangulation Supports Agricultural Spread of the Transeurasian Languages. **Nature**, London, v. 599, p. 616-621, 10 November 2021.
- ROBBEETS, M.; BISANG, W. When Paradigms Change. In: ROBBEETS, M.; BISANG, W. **Paradigm Change: In The Transeurasian languages and beyond**. Amsterdam, NL: John Benjamins Publishing Co., 2014. p. 1-22.
- SETH, M. J. **A History of Korea: From Antiquity to the Present**. Plymouth, UK: Rowman & Littlefield Publishers, 2011.
- SHIBATANI, M. **The Languages of Japan**. New York, NY: Cambridge University Press, 1990.
- SHIMOJI, M. Northern Ryukyuan. In: TRANTER, N. **The Languages of Japan and Korea**. New York, NY: Routledge, 2012. p. 351-380.
- SOHN, H.-M. **The Korean Language**. New York, NY: Cambridge University Press, 1999.
- SUN, C. **Chinese: A Linguistic Introduction**. Cambridge, UK: Cambridge University Press, 2006.
- TANAKA DE LIRA, K.; TANAKA DE LIRA, M. De Aomori ao Shuri: Excetos dos Dialectos da Língua Japonesa. **Estudos Japoneses**, São Paulo, v. 36, p. 71-93, 2016.
- TANAKA DE LIRA, M. Classes Adjetivais no Nordeste Asiático: Revendo as Línguas Japonesa e Coreana. **Hon No Mushi**, Manaus, AM, v. 5, p. 50-83, ago. 2020. ISSN 8.
- TANAKA DE LIRA, M. O Nordeste Asiático como Área de Convergência Linguística: a Língua Japonesa em seu Contexto Regional. In: JOKO, A. T., et al. **Diálogo Linguístico: Ocidente e Oriente**. Brasília: Editora Universidade de Brasília, 2021. p. 315-336.
- TANAKA DE LIRA, M. Formação Linguística do Nordeste Asiático: As Múltiplas Origens das Famílias Coreânica e Japônica. **Hon no Mushi**, Manaus, No Prelo.
- THORPE, M. L. **Ryūkyūan Language History**. Los Angeles: University of Southern California (Dissertation), 1983.
- TŌKEIKYOKU. 統計でみる都道府県のすがた | **Statistical Observations of Prefectures**. Tokyo, JP: Statistics Bureau, Ministry of Internal Affairs and Communications of Japan, 2023.

- UNGER, J. M. When was Korean First Poken in Southeastern Korea? **Journal of Inner and East Asian Studies**, Seoul, v. 2, n. 2, p. 87-107, December 2005.
- UNIVERSITY OF VIRGINIA LIBRARY ELECTRONIC TEXT CENTER. Man'yōshū. **Man'yōshū**, 759. Disponível em: <<http://jti.lib.virginia.edu/japanese/manyoshu/AnoMany.html>>. Acesso em: 11 15 2022.
- VOVIN, A. Koguryō and Paekche: Different Languages or Dialects of Old Korean? **Journal of Inner and East Asian Studies**, Seoul, v. 2, n. 2, p. 107-140, December 2005.
- VOVIN, A. **Koreo-Japonica: A Re-Evaluation of a Common Genetic Origin**. Honolulu, HI: University of Hawai'i Press, 2010.
- VOVIN, A. **A Descriptive and Comparative Grammar of Western Old Japanese (Revised, Updated and Enlarged)**. 2ª. ed. Boston, MA: Brill, v. 1, 2020.
- WADE, N. Finding on Dialects Casts New Light on the Origins of the Japanese People. **The New York Times**, New York, NY, 04 maio 2011. A-16.
- WARDHAUGH, R. **An Introduction to Sociolinguistics**. 5ª. ed. Malden, MA: Blackwell Publishing, 2006.
- WEISS, M. The Comparative Method. In: BOWERN, C.; EVANS, B. **The Routledge Handbook of Historical Linguistics**. New York, NY: [s.n.], 2015. p. 127-145.
- WESTNEY, D. E. The Emulation of Western Organizations in Meiji Japan: the case of the Paris Prefecture of Police and the Keishi-chō. In: KORNICKI, P. **Meiji Japan: Political, Economic and Social History 1868-1912 (Volume III: The Mature Meiji State)**. London, UK: Routledge, 1998. p. 107-137.
- WHITMAN, J. The Relationship between Japanese and Korean. In: TRANTER, N. **The Languages of Japan and Korea (Routledge Language Family Series)**. New York, NY: Routledge, 2012. p. 24-38.
- YANG, C.; YANG, S.; O'GRADY, W. **Jejueo: The Language of Korea's Jeju Island**. Honolulu, HI: University of Hawai'i Press, 2020.
- YEON, J. Korean Dialects: a General Survey. In: TRANTER, N. **The Languages of Japan and Korea (Routledge Language Family Series)**. New York, NY: Routledge, 2012. p. 168-186.
- ZHENGZHANG, S.; ZHENG, W. Wu Dialect. In: SUN, C.; WANG, W. S.-Y.; TSAI, Y. **The Oxford Handbook of Chinese Linguistics**. New York, NY: Oxford University Press, 2015. p. 189-202.

*Recebido em 07 de fevereiro de 2023  
Aprovado em 22 de agosto de 2023*