

---

# Computational Linguistics-1

---

Summer 2021, IIIT Hyderabad

---

## Graded Exercise 2

---

### Original text (10 sentences)

---

The State government would impose a strict lockdown in the State from June 5 to June 9 to curb COVID-19 transmission.

Chief Minister Pinarayi Vijayan told a top-level pandemic review meeting on Thursday that the government would allow only essential services during the period.

The relaxation on curbs on life, mobility, retail and commerce would end at 7 PM on June 4th.

Only shops selling essential supplies, medicine, construction material, packaging and raw material for industrial production could open for business.

Stores dealing in scrap could conduct business till 7 PM on June 4th.

Government, quasi-government, and public sector institutions, corporations, and various commissions would remain closed.

They could open with less than 50% of the workforce from June 10th.

People entering Kerala would require a COVID negative certificate.

The State has braced itself for a third wave.

As a first step , the administration would fully inoculate guest workers and other vulnerable sections.

---

### Tagged (using CLAWS7)

The\_AT State\_NN1 government\_NN1 would\_VM impose\_VVI a\_AT1 strict\_JJ lockdown\_NN1 in\_II the\_AT State\_NN1 from\_II June\_NPM1 5\_MC to\_II June\_NPM1 9\_MC to\_TO curb\_VVI COVID-19\_MC transmission\_NN1 .\_.

Chief\_JJ Minister\_NN1 Pinarayi\_NP1 Vijayan\_NP1 told\_VVD a\_AT1 top-level\_JJ pandemic\_JJ review\_NN1 meeting\_VVG on\_II Thursday\_NPD1 that\_CST the\_AT government\_NN1 would\_VM allow\_VVI only\_RR essential\_JJ services\_NN2 during\_II the\_AT period\_NN1 .\_.

The\_AT relaxation\_NN1 on\_II curbs\_NN2 on\_II life\_NN1 ,\_, mobility\_NN1 ,\_, retail\_JJ and\_CC commerce\_NN1 would\_VM end\_VVI at\_II 7\_MC PM\_NP1 on\_II June\_NPM1 4th\_MD .\_.

Only\_JJ shops\_NN2 selling\_VVG essential\_JJ supplies\_NN2 ,\_, medicine\_NN1 ,\_, construction\_NN1 material\_NN1 ,\_, packaging\_NN1 and\_CC raw\_JJ material\_NN1 for\_IF industrial\_JJ production\_NN1 could\_VM open\_JJ for\_IF business\_NN1 .\_.

Stories\_NN2 dealing\_VVG in\_II scrap\_NN1 could\_VM conduct\_VVI business\_NN1 till\_II 7\_MC PM\_NP1 on\_II June\_NPM1 4th\_MD \_.\_.

Government\_NN1 ,\_, quasi-government\_NN1 ,\_, and\_CC public\_JJ sector\_NN1 institutions\_NN2 ,\_, corporations\_NN2 ,\_, and\_CC various\_JJ commissions\_NN2 would\_VM remain\_VVI closed\_JJ \_.\_.

They\_PPHS2 could\_VM open\_VVI with\_IW less\_DAR than\_CSN 50%\_NNU of\_IO the\_AT workforce\_NN1 from\_II June\_NPM1 10th\_MD \_.\_.

People\_NN entering\_VVG Kerala\_NP1 would\_VM require\_VVI a\_AT1 COVID\_JJ negative\_JJ certificate\_NN1 \_.\_.

The\_AT State\_NN1 has\_VHZ braced\_VVN itself\_PPX1 for\_IF a\_AT1 third\_MD wave\_NN1 \_.\_.

As\_II a\_AT1 first\_MD step\_NN1 ,\_, the\_AT administration\_NN1 would\_VM fully\_RR inoculate\_VVI guest\_NN1 workers\_NN2 and\_CC other\_JJ vulnerable\_JJ sections\_NN2 \_.\_.

---

## Translated (using Google Translate)

COVID-19 പ്രക്ഷേപണം തടയുന്നതിന് സംസ്ഥാന സർക്കാർ ജൂൺ 5 മുതൽ ജൂൺ 9 വരെ കർശനമായി ലോക്ക്ഡൗൺ ഏർപ്പെടുത്തും.

COVID-19 prakṣeṇaṁ taṭayunnatin' samstha:na sarkka:r ju:ṇ 5 mutal ju:ṇ 9 vare karfanamay:ilokkḍaṇ eṛpeṭuttum.

ഈ കാലയളവിൽ അവശ്യ സേവനങ്ങൾ മാത്രമേ ഭരണകൂടം അനുവദിക്കുകയുള്ളൂവെന്ന് മുഖ്യമന്ത്രി പിണറായി വിജയൻ വ്യാഴാഴ്ച ഉന്നതതല പാൻഡെമിക് അവലോകന യോഗത്തിൽ പറഞ്ഞു.

i: kaḷayaḷavil avayya se:vananṇaḷ ma:trame: bharaṇaku:ṭam anuvadikkukayulḷu:venn' mukhyamantri piṇaRayi vijayan vya:zha:zhca unnatatala pa:ṇḍemik avalo:kana yo:gattil paṇṇu.

ജീവിതം, മൊബിലിറ്റി, റീടെയിൽ, വാണിജ്യം എന്നിവയ്ക്കുള്ള ഇളവ് ജൂൺ 4 ന് വൈകുന്നേരം 7 മണിക്ക് അവസാനിക്കും.

ji:vitam, mobiliRRi, Riṭṭeyil, vaṇiyyam ennivaykkuḷḷa ilav' ju:ṇ na:lin' vaikunne:ram 7 maṇikk' avasa:nikkum.

അവശ്യസാധനങ്ങൾ, മരുന്ന്, നിർമ്മാണ സാമഗ്രികൾ, പാക്കേജിംഗ്, വ്യാവസായിക ഉൽപ്പാദനത്തിനുള്ള അസംസ്കൃത വസ്തുക്കൾ എന്നിവ വിൽക്കുന്ന കടകൾക്ക് മാത്രമേ ബിസിനസിന് തുറക്കാൻ കഴിയൂ.

avayya:dhanaṇaḷ, marunn', nirmma:ṇa sa:magrikaḷ, pa:kke:jimg', vya:vasa:yika ulpa:danattinulḷa asamskrta vastukkaḷ enniva vilkkunna kaṭakaḷkk' ma:trame: bisinasin' turakka:n kazhiyū.

സ്ക്രാപ്പ് കൈകാര്യം ചെയ്യുന്ന സ്റ്റോറുകൾക്ക് ജൂൺ 4 ന് വൈകുന്നേരം 7 മണി വരെ ബിസിനസ്സ് നടത്താം. skra:pp' kaika:ryam ceyyunna sRRo:Rukaḷkk' ju:ṇ na:lin' vaikunne:ram 7 maṇi vare bisinass' naṭatta:m

സർക്കാർ, അർദ്ധ സർക്കാർ, പൊതുമേഖലാ സ്ഥാപനങ്ങൾ, കോർപ്പറേഷനുകൾ, വിവിധ കമ്മീഷനുകൾ എന്നിവ അടച്ചിരിക്കും.

sarkka:r, arddha sarkka:r, potume:khalā: stha:pananṇaḷ, ko:rppaRe:ṣanukaḷ, vividha kammi:ṣanukaḷ enniva aṭaccirikkum.

ജൂൺ 10 മുതൽ 50% ൽ താഴെ തൊഴിലാളികളുമായി അവർക്ക് തുറക്കാൻ കഴിയും.

ju:ṇ 10 mutal 50% il tazhe tozhila:ḷikaḷuma:y: avarkk' turakka:n kazhiyum.

കേരളത്തിലേക്ക് പ്രവേശിക്കുന്ന ആളുകൾക്ക് COVID നെഗറ്റീവ് സർട്ടിഫിക്കറ്റ് ആവശ്യമാണ്.

ke:raḷattile:kk' prave:fikkunna aḷukaḷkk' COVID negaRRi:v' sarṭṭiphikkaRR' a:vayama:ṇ'.

മൂന്നാമത്തെ തരംഗത്തിന് സംസ്ഥാനം സ്വയം ശ്രമിച്ചു.  
mu:nnamatte taramgattin' samstha:nam svayam framiccu.

ആദ്യ ഘട്ടമെന്ന നിലയിൽ, ഗസ്റ്റ് വർക്കർമാരെയും മറ്റ് ദുർബല വിഭാഗങ്ങളെയും ഭരണകൂടം പൂർണ്ണമായും കുത്തിവയ്ക്കും.

a:dya ghaṭṭamenna nilayil, gasRR' varkkarma:reyum maRR' durbala vibha:gaṇṇaḷeyum  
bharanakuṭam pu:ṇṇama:yum kuttivaykkum.

---

## Post-edited

COVID-19 പ്രക്ഷേപണം തടയുന്നതിന് സംസ്ഥാന സർക്കാർ ജൂൺ 5 മുതൽ ജൂൺ 9 വരെ കർശനമായ ലോക്ക്ഡൗൺ ഏർപ്പെടുത്തും.

COVID-19 prakṣe:paṇam taṭayunnatin' samstha:na sarkka:r ju:ṇ 5 mutal ju:ṇ 9 vare karṇanama:ya  
lo:kdauṇ e:rpeṭuttum.

ഈ കാലയളവിൽ അവശ്യ സേവനങ്ങൾ മാത്രമേ ഭരണകൂടം അനുവദിക്കുകയുള്ളൂവെന്ന് മുഖ്യമന്ത്രി പിണറായി വിജയൻ വ്യാഴാഴ്ച ഉന്നതതല പാൻഡെമിക് അവലോകന യോഗത്തിൽ പറഞ്ഞു.

i: ka:layaḷavil avafya se:vananṇaḷ ma:trame: bharanakuṭam anuvadikkukayul[u:venn' mukhyamantri  
piṇaRayi vijayan vyaz:hazhca unnatatala pa:ṇḍemik avalo:kana yo:gattil paRaṇṇu.

ജീവിതം, മൊബിലിറ്റി, റീട്ടെയിൽ, വാണിജ്യം എന്നിവയ്ക്കുള്ള ഇളവ് ജൂൺ 4 ന് വൈകുന്നേരം 7 മണിക്ക് അവസാനിക്കും.

ji:vitam, mobiliRRi, Ri:ṭṭeyil, vaṇijyam ennivaykkul[a ilav' ju:ṇ na:lin' vaikunne:ram 7 maṇikk'  
avasaniikkum.

അവശ്യസാധനങ്ങൾ, മരുന്നുകൾ, നിർമ്മാണ സാമഗ്രികൾ, പാക്കേജിംഗ്, വ്യാവസായിക

ഉൽപാദനത്തിനുള്ള അസംസ്കൃത വസ്തുക്കൾ എന്നിവ വിൽക്കുന്ന കടകൾക്ക് മാത്രമേ ബിസിനസിന് തുറക്കാൻ കഴിയൂ.

avafyasa:dhananṇaḷ, marunnukaḷ, nirmma:ṇa sa:magrikaḷ, pa:kke:jimg', vya:vasa:yika  
ulpa:danattinul[a asamskrta vastukkaḷ enniva vilkkunna kaṭakaḷkk' ma:trame: bisinasin' turakka:n  
kazhiyu:.

സ്ക്രാപ്പ് കൈകാര്യം ചെയ്യുന്ന സ്റ്റോറുകൾക്ക് ജൂൺ 4 ന് വൈകുന്നേരം 7 മണി വരെ ബിസിനസ്സ് നടത്താം.  
skra:pp' kaika:ryam ceyyunna sRRo:Rukaḷkk' ju:ṇ na:lin' vaikunne:ram 7 maṇi vare bisinass'  
naṭatta:m.

സർക്കാരും അർദ്ധ സർക്കാരും, പൊതുമേഖലയും സ്ഥാപനങ്ങൾ, കോർപ്പറേഷനുകൾ, വിവിധ കമ്മീഷനുകൾ എന്നിവ അടച്ചിടിക്കും.

sarkka:Rum, arddha sarkka:Rum potume:khalayum stha:pananṇaḷ, ko:rppaRe:ṣanukaḷ, vividha  
kammi:ṣanukaḷ enniva aṭaccirikkum.

ജൂൺ 10 മുതൽ 50% ൽ താഴെ തൊഴിലാളികളുമായി അവയ്ക്ക് തുറക്കാൻ കഴിയും.

ju:ṇ 10 mutal 50% il tazhe tozhila:ḷikaḷuma:yi avaykk' turakka:n kazhiyum.

കേരളത്തിലേക്ക് പ്രവേശിക്കുന്ന ആളുകൾക്ക് COVID നെഗറ്റീവ് സർട്ടിഫിക്കറ്റ് ആവശ്യമാണ്.

ke:raḷattile:kk' prave:fikkunna a:ḷukaḷkk' COVID negaRRi:v' sarṭṭiphikkaRR' a:avafyama:ṇ'.

മൂന്നാമത്തെ തരംഗത്തിന് സംസ്ഥാനം സ്വയം ശ്രമിച്ചു.

mu:nnamatte taramgattin' samstha:nam svayam framiccu.

ആദ്യ ഘട്ടമെന്ന നിലയിൽ, ഗസ്റ്റ് വർക്കർമാരെയും മറ്റ് ദുർബല വിഭാഗങ്ങളെയും ഭരണകൂടം പൂർണ്ണമായും കുത്തിവയ്ക്കും.

a:dya ghaṭṭamenna nilayil, gasRR' varkkarma:reyum maRR' durbala vibha:gaṇṇaḷeyum  
bharaṇaku:ṭam pu:rṇama:yum kuttivaykkum.

## Tagged

COVID-19\_NN1 പ്രക്ഷേപണം\_NN1 തടയുന്നതിന്\_VVGN-IF സംസ്ഥാന\_JJ സർക്കാർ\_NN1 ജൂൺ\_NPM  
5\_MC മുതൽ\_II ജൂൺ\_NPM 9\_MC വരെ\_II കർശനമായ\_JJ ലോക്ക്ഡൗൺ\_NN1 ഏർപ്പെടുത്തും\_VVF.  
COVID-19\_NN1 prakṣe:paṇam\_NN1 taṭayunnatin'\_VVGN-IF samstha:na\_JJ sarkka:r\_NN1  
ju:ṇ\_NPM 5\_MC mutal\_II ju:ṇ\_NPM 9\_MC vare\_II karṇanama:ya\_JJ lo:kḍauṇ\_NN1  
e:rpeṭuttum\_VVF.

ഈ\_DDP കാലയളവിൽ\_NN1-II അവശ്യ\_JJ സേവനങ്ങൾ\_NN2 മാത്രമേ\_RR ഭരണകൂടം\_NN1  
അനുവദിക്കുകയുള്ളൂ\_VVF-O എന്ന്\_CST മുഖ്യമന്ത്രി\_NN1 പിണറായി\_NP1 വിജയൻ\_NP1  
വ്യാഴാഴ്ച\_NPD1 ഉന്നതതല\_JJ പാൻഡെമിക്\_NN1 അവലോകന\_NN1 യോഗത്തിൽ\_NN-II  
പറഞ്ഞു\_VVD.

i:DDP ka:layaḷavil\_NN1-II avaḥya\_JJ se:vanarṇaḷ\_NN2 maṭrame:RR bharaṇaku:ṭam\_NN1  
anuvadikkukayulḷu:VVF-O enn'\_CST mukhyamantri\_NN1 piṇaRayi\_NP1 vijayan\_NP1  
vyazha:zhca\_NPD1 unnatatala\_JJ pa:ṇḍemik\_NN1 avalo:kana\_NN1 yo:gattil\_NN-II paRaṇṇu\_VVD.

ജീവിതം\_NN1, മൊബിലിറ്റി\_NN1, റീട്ടെയിൽ\_NN1, വാണിജ്യം\_NN1 എന്നിവയ്ക്കുള്ള\_PH2NP-IF-EXJ  
ഇളവ്\_NN1 ജൂൺ\_NPM 4\_MC ന്\_II വൈകുന്നേരം\_NPD1 7\_MC മണിക്ക്\_NN1-II അവസാനിക്കും\_VVF.  
ji:vitam\_NN1, mobiliRRi\_NN1, Ri:ṭṭeyil\_NN1, vaṇijyam\_NN1 ennivaykkuḷḷa\_PH2NP-IF-EXJ  
iḷav'\_NN1 ju:ṇ\_NPM na:lin'\_MC-II vaikunne:ram\_NPD1 7\_MC maṇikk'\_NN1-II avasa:nikkum\_VVF.

അവശ്യസാധനങ്ങൾ\_NN2, മരുന്നുകൾ\_NN2, നിർമ്മാണ\_JJ സാമഗ്രികൾ\_NN2, പാക്കേജിംഗ്\_NN1,  
വ്യാവസായിക\_JJ ഉൽപാദനത്തിനുള്ള\_NN1-IF-EXJ അസംസ്കൃത\_JJ വസ്തുക്കൾ\_NN2  
എന്നിവ\_PH2NP വിൽക്കുന്ന\_VVGJ കടകൾക്ക്\_NN2-IF മാത്രമേ\_RR ബിസിനസിന്\_NN1-IF  
തുറക്കാൻ\_VDI കഴിയൂ\_VVF-O.

avaḥyasa:dhanarṇaḷ\_NN2, marunnukaḷ\_NN1, nirmma:ṇa\_JJ sa:magrikaḷ\_NN2, pa:kke:jimg'\_NN1,  
vya:asa:yika\_JJ ulpa:danattinulḷa\_NN1-IF-EXJ asamskṛta\_JJ vastukkaḷ\_NN2 enniva\_PH2NP  
vilkkunna\_VVGJ kaṭakaḷkk'\_NN2-IF maṭrame:RR bisinasin'\_NN1-IF turakka:n\_VDI kazhiyu:VVF-O.

സ്ക്രാപ്പ്\_NN1 കൈകാര്യം\_NN1 ചെയ്യുന്ന\_VVGJ സ്റ്റോറുകൾക്ക്\_NN2-IF ജൂൺ\_NPM 4\_MC ന്-IF  
വൈകുന്നേരം\_NPD1 7\_MC മണി\_NPD1 വരെ\_II ബിസിനസ്സ്\_NN1 നടത്താം\_VVP.  
skra:pp'\_NN1 kaika:ryam\_NN1 ceyyunna\_VVGJ sRRo:Rukaḷkk'\_NN2-IF ju:ṇ\_NPM na:lin'\_MC-IF  
vaikunne:ram\_NPD1 7\_MC maṇi\_NPD1 vare\_II bisinass'\_NN1 naṭatta:m\_VVP.

സർക്കാരും\_NN1-CC അർദ്ധ\_JJ സർക്കാരും\_NN1-CC, പൊതുമേഖലയും\_NN1-CC സ്ഥാപനങ്ങൾ\_NN2,  
കോർപ്പറേഷനുകൾ\_NN2, വിവിധ\_JJ കമ്മീഷനുകൾ\_NN2 എന്നിവ\_PH2NP അടച്ചിരിക്കും\_VVF.  
sarkka:Rum\_NN1-CC, arddha\_JJ sarkka:Rum\_NN1-CC potume:khalayum\_NN1-CC  
stha:panarṇaḷ\_NN2, ko:rppaRe:ṣanukaḷ\_NN2, vividha\_JJ kammi:ṣanukaḷ\_NN2 enniva\_PH2NP  
aṭaccirikkum\_VVF.

ജൂൺ\_NPM 10\_MC മുതൽ\_II 50%\_DB ൽ\_II താഴെ\_RL തൊഴിലാളികളുമായി\_NN2-II  
അവയ്ക്ക്\_PH2NR തുറക്കാൻ\_VDI കഴിയും\_VVF.  
ju:ṇ\_NPM 10\_MC mutal\_II 50%\_DB il-II ta:zhe\_RL tozhila:ḷikaḷuma:yi\_NN2-II avaykk'\_PH2NR  
turakka:n\_VDI kazhiyum\_VVF.

കേരളത്തിലേക്ക്\_NP1-II പ്രവേശിക്കുന്ന\_VVGJ ആളുകൾക്ക്\_NN2-IF COVID\_NN1 നെഗറ്റീവ്\_JJ  
സർട്ടിഫിക്കറ്റ്\_NN1 ആവശ്യമാണ്\_JJ-VBZ.  
ke:raḷattile:kk'\_NP1-II prave:fikkunna\_VVGJ aḷukaḷkk'\_NN2-IF COVID\_NN1 negaRRi:v'\_JJ  
sarṭṭiphikkaRR'\_NN1 a:vaḥyama:ṇ'\_JJ-VBZ.

മൂന്നാമത്തെ\_MD തരംഗത്തിന്\_NN1-IF സംസ്ഥാനം\_NN1 സ്വയം\_PPX ശ്രമിച്ചു\_VVD.  
mu:nnamatte\_MD taramgattin'\_NN1-IF samstha:nam\_NN1 svayam\_PPX framiccu\_VVD.

ആദ്യ\_JJ ഘട്ടം\_NN1 എന്ന\_II നിലയിൽ\_II, ഗസ്റ്റ്\_NN1 വർക്കർമാരെയും\_NN2-II-CC മറ്റ്\_JJ ദുർബല\_JJ  
വിഭാഗങ്ങളെയും\_NN2-II-CC ഭരണകൂടം\_NN1 പൂർണ്ണമായും\_RR കുത്തിവയ്ക്കും\_VVF.  
a:dya\_JJ ghaṭṭamenna\_NN1 nilayil\_II, gasRR'\_NN1 varkkarma:reyum\_NN2-II-CC maRR'\_JJ  
durbala\_JJ vibha:gaṇṇa[eyum\_NN2-II-CC bharaṇakuṭam\_NN1 pu:rṇama:yum\_RR  
kuttivaykkum\_VVF.

---

## Extensions to CLAWS7 Tagset

Verbs:

VVF: simple future/present habitual tense (suffix -um “-ഉം”)

VVC: present continuous (suffix -unnu “-ഉന്നു”)

VVP: potential (suffix -a:m “-ആം”)

VVGn: present gerund (noun)

VVGJ: present gerund (adjective)

EXJ: existential gerund (adjective)

-O: “only” (suffix -u: “-ഊ”; accompanies adverb ma:trame: “മാത്രമേ”)

Conjunctions:

CST: quotative particle enn' “എന്ന്” (used as “that” in English)

Pronouns:

PH2NP: neuter 3rd person pronoun plural proximate

PH2NR: neuter 3rd person impersonal pronoun plural remote

PPX: reflexive pronoun (svayam “സ്വയം”)

Note: Tags glossed as “prepositions” in CLAWS7 are treated as postpositional tags here.

---

## Notes on Transliteration

Velar, dental and bilabial non-nasal stops are as in IAST and common transliteration.

The velar nasal is ŋ; the dental nasal is ñ; and the bilabial nasal is m.

The palatal consonants are c, ch, j, jh, ṇ.

The retroflex consonants are ṭ, ṭh, ḍ, ḍh, ṇ.

The semivowels common to Malayalam and Devanagari are transliterated the common way (y, r, l, v).

The sibilants are respectively ṣ (retroflex: Devanagari ष and Telugu ష), ś (palatal: Devanagari श and Telugu శ), s and h.

The retroflex lateral approximant is ɭ (Devanagari ळ and Telugu ఱ) and the voiced retroflex approximant is zh.

The alveolar trill is R (as opposed to the alveolar tap which is r).

RR is a direct transliteration of the letter for R, doubled; it is pronounced as an alveolar voiceless stop (distinct from the retroflex and dental voiceless stops).

m occurring at the end of words or immediately before consonants represents the equivalent of Devanagari ॠ (अं) and Telugu ౠ (అం).

r occurring between two consonants represents the syllabic r (Devanagari ॠ and Telugu ౠ).

' (apostrophe) at the end of words represents the close central unrounded vowel ɨ. In the script it is represented by either the equivalent of the halant or the symbol for 'u'. It does not have an independent vowel sign (without a consonant).

Long vowels are marked by the : symbol.