

日本語LGRの文字範囲 と異体字定義について

1. 日本語LGRの登録可能文字範囲

- 考えられる選択肢

- a. 常用漢字・人名用漢字の範囲

日本国内の地名や歴史的名称を表現するには不十分である。
例えば奈良県橿原市の「橿」はこの範囲に含まれない

- b. JIS X 0208:2012の第一水準・第二水準の範囲

常用漢字(JIS X 0213:2012の第三水準に含まれる4文字を除く)・
人名用漢字を包含し、かつ日本国内の地名や歴史的名称を表
記でき、JPドメイン名でも有用性が確認されている(文字の多寡
について苦情や要望がない)

- c. JIS X 0213:2012の第三水準・第四水準までの範囲

平均的な日本人が通常使用しておらず、また理解していない(読
み書きできない)文字を入れるのはリスクが大きい

- d. IICOREの範囲

JIS X 0203:2012と同様

2. 日本語LGRの異体字定義

2. 考えられる選択肢

- a. 中国語LGRで定義される異体字を日本語LGRでも異体字として定義する

日本語LGRが異体字を定義しない場合、中国語LGRで定義される異体字はRootLGRですべてブロックされる扱いとなる(どちらも使いたいという要求に応えられない)ため、それを避ける(登録可能となるようにする)ためには日本語LGRでも受身的に異体字として定義する必要がある

- b. 日本語LGRでも積極的(独自)に異体字を定義する

ドメイン名というコンテキストに適切な異体字を適切な根拠の下に定義しオーソリティを得る

本資料で説明するもの

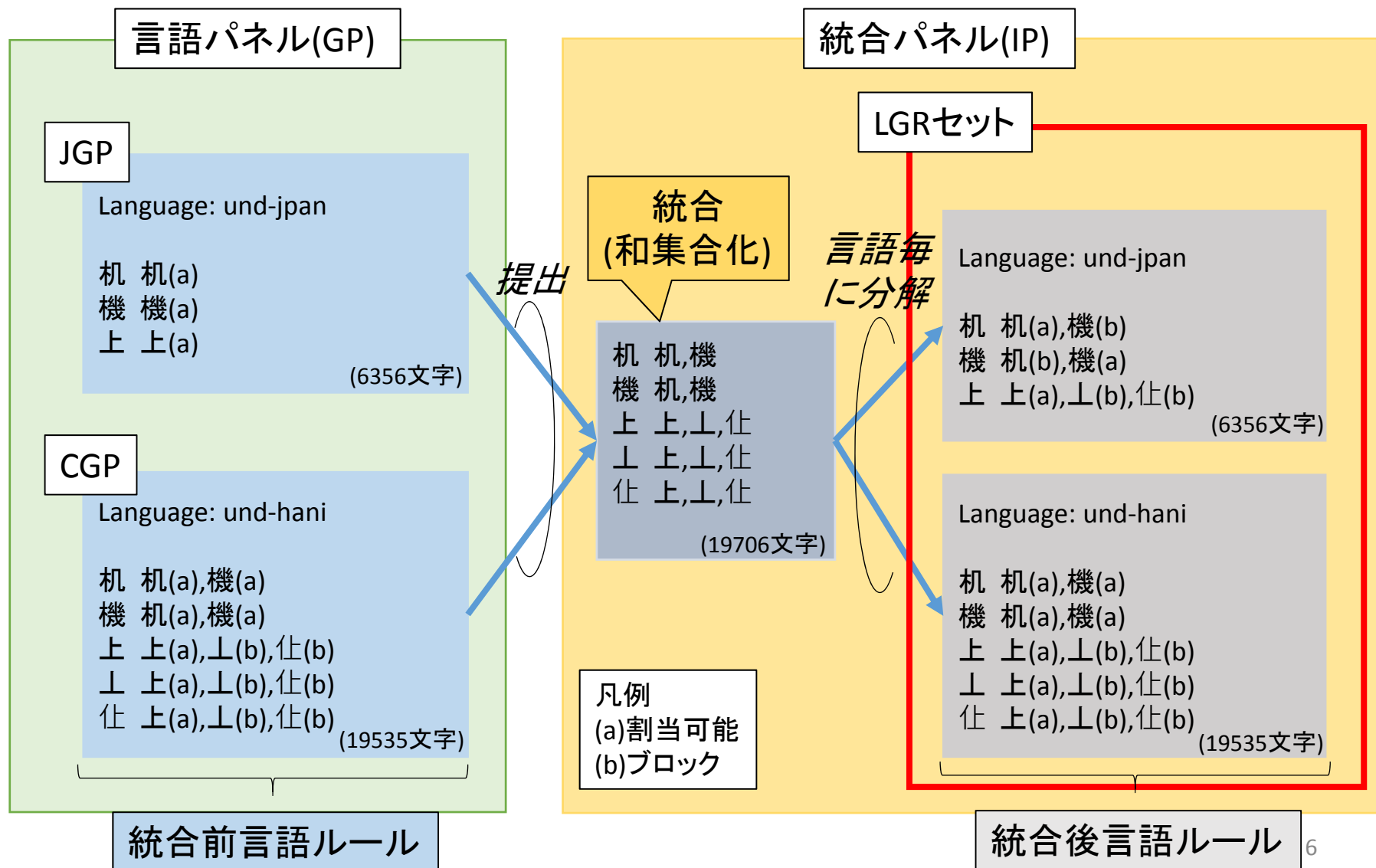
- JGPとCGPが作成する言語LGRのイメージと、IPが統合したRootLGRのイメージを示す
- 文字の範囲は1-b「JIS X 0208:2012の第一水準・第二水準の範囲」を仮定する
- 日本語LGRでの異体字定義は、以下のパターンで例を示す
 - JGPは異体字を定義しない
 - JGPは受身的に異体字を定義する(2-a)
 - JGPは積極的に異体字を定義する(2-b)

機上についての考察

- 前提

1. 中国語では「機」と「机」は異体字である
2. 中国語では「上」と「上」と「仕」は異体字である

例1: JGPは異体字を定義しない



例1: 文字列生成例

【注】

混在可能: (a)と(a)

混在不可能: (a)と(b)、(b)と(b)

Language: und-jpan

申請文字列: 機上

割当可能: 機上

ブロック: 機上, 機上, 机上, 机上, 机上

Language: und-hani

申請文字列: 機上

割当可能: 機上, 机上

ブロック: 機上, 機上, 机上, 机上

Language: und-jpan

申請文字列: 机上

割当可能: 机上

ブロック: 机上, 机上, 機上, 機上, 機上

Language: und-hani

申請文字列: 机上

割当可能: 机上, 機上

ブロック: 机上, 机上, 機上, 機上

Language: und-jpan

申請文字列: 机機

割当可能: 机機

ブロック: 机机, 機机, 機機

Language: und-hani

申請文字列: 机上

割当可能: 机上, 機上

ブロック: 机上, 机上, 機上, 機上

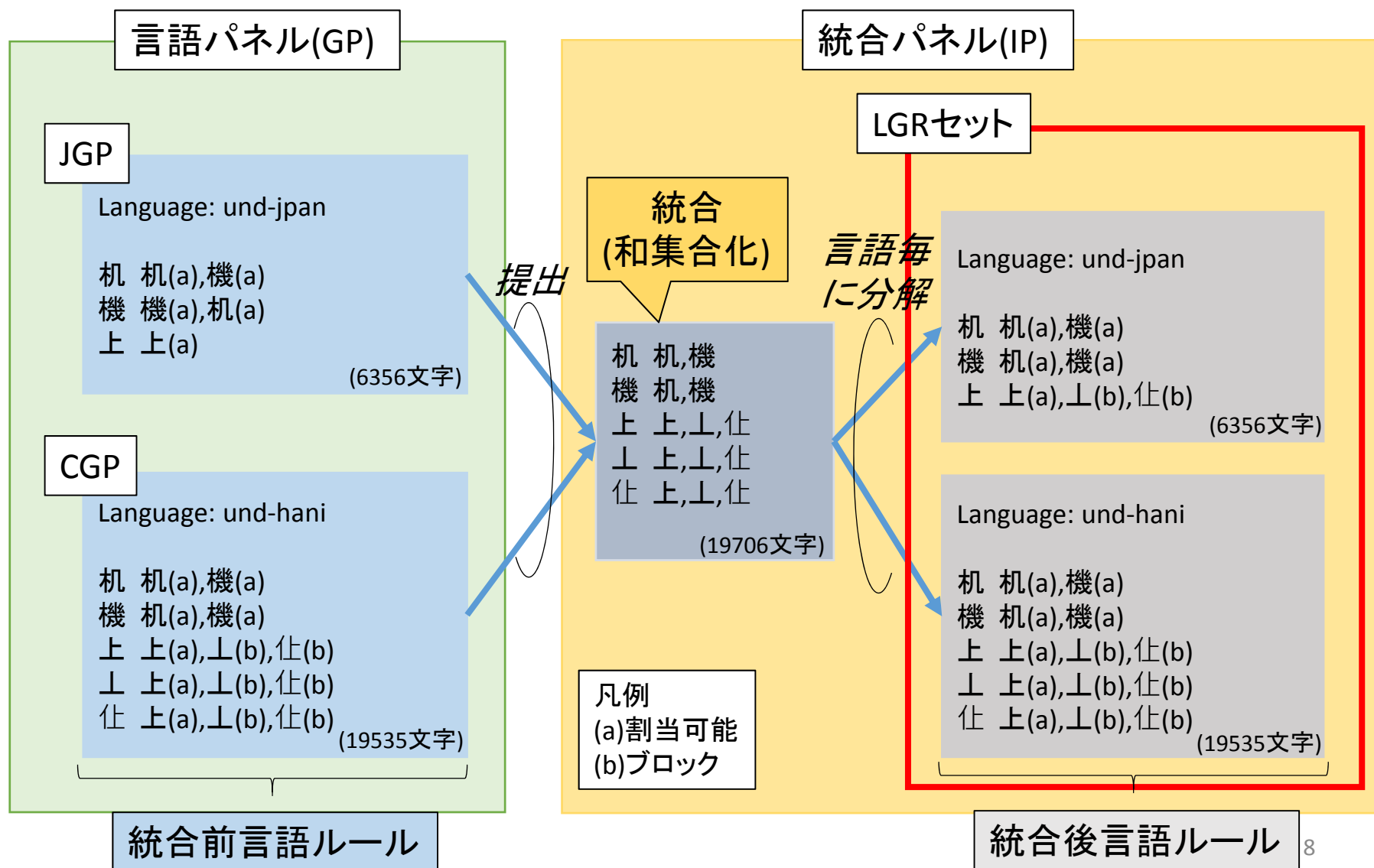
Language: und-hani

申請文字列: 上上

割当可能: 上上

ブロック: 上上, 上上, 上上, 上上, 上上, 上上, 上上, 上上

例2: JGPは受身的に異体字を定義する



例2: 文字列生成例

【注】

混在可能: (a)と(a)

混在不可能: (a)と(b)、(b)と(b)

Language: und-jpan
申請文字列: 机上
割当可能: 机上, 机上
ブロック: 机上, 机上, 机上, 机上

Language: und-hani
申請文字列: 机上
割当可能: 机上, 机上
ブロック: 机上, 机上, 机上, 机上

Language: und-jpan
申請文字列: 机上
割当可能: 机上, 机上
ブロック: 机上, 机上, 机上, 机上

Language: und-hani
申請文字列: 机上
割当可能: 机上, 机上
ブロック: 机上, 机上, 机上, 机上

Language: und-jpan
申請文字列: 机機
割当可能: 机机, 机機, 机機, 机機
ブロック:

Language: und-hani
申請文字列: 机上
割当可能: 机上, 机上
ブロック: 机上, 机上, 机上, 机上

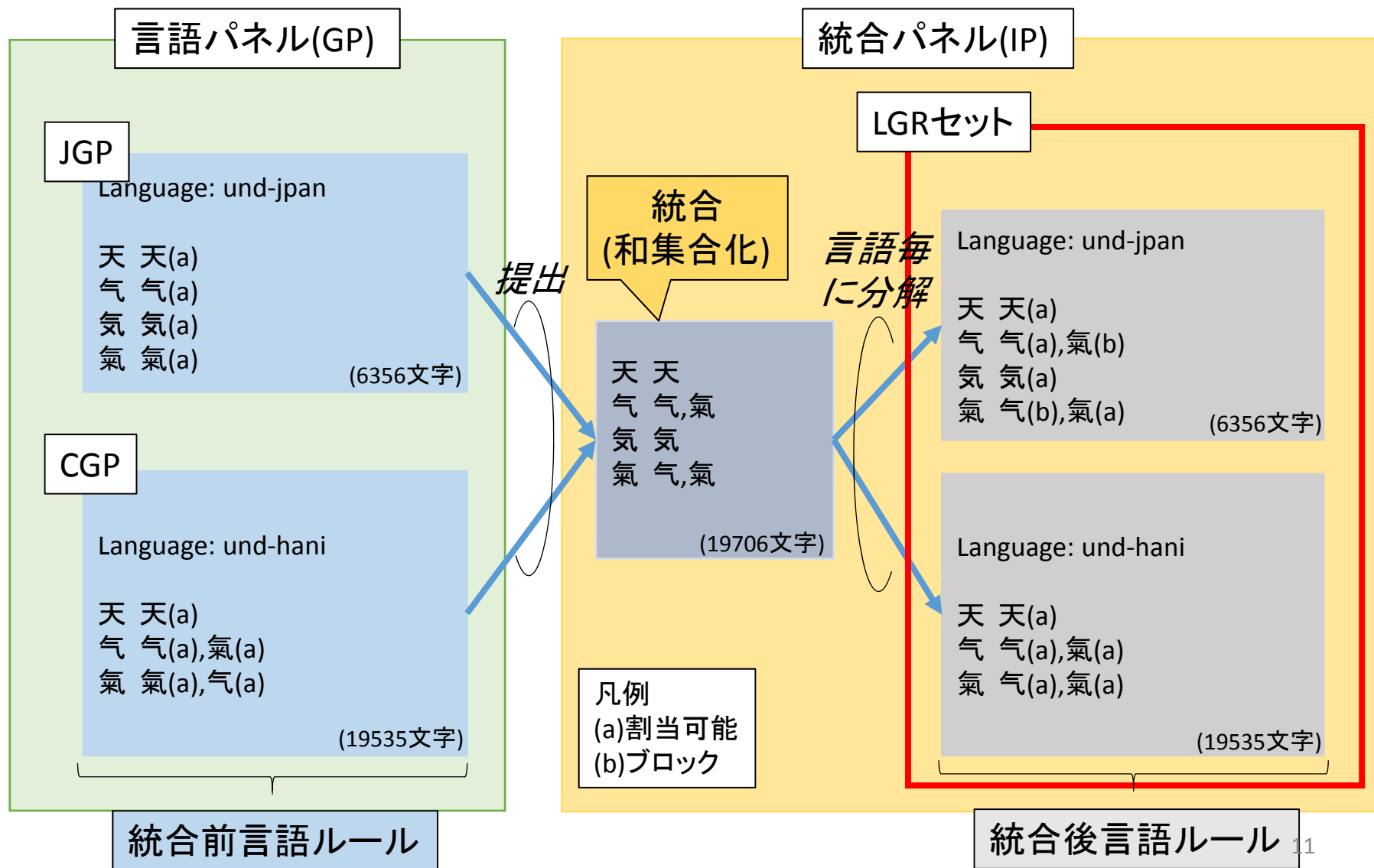
Language: und-hani
申請文字列: 上上
割当可能: 上上
ブロック: 上上, 上上, 上上, 上上, 上上, 上上, 上上, 上上

天気についての考察

- 前提

1. 中国語では「气」と「氣」は異体字である
2. 日本語では「気」と「氣」を異体字とする文献がある
3. 「気」は日本語の国字である

例3: JGPは異体字を定義しない



例3: 文字列生成例

【注】

混在可能: (a)と(a)

混在不可能: (a)と(b)、(b)と(b)

Language: und-jpan
申請文字列: 天氣
割当可能: 天氣
ブロック:

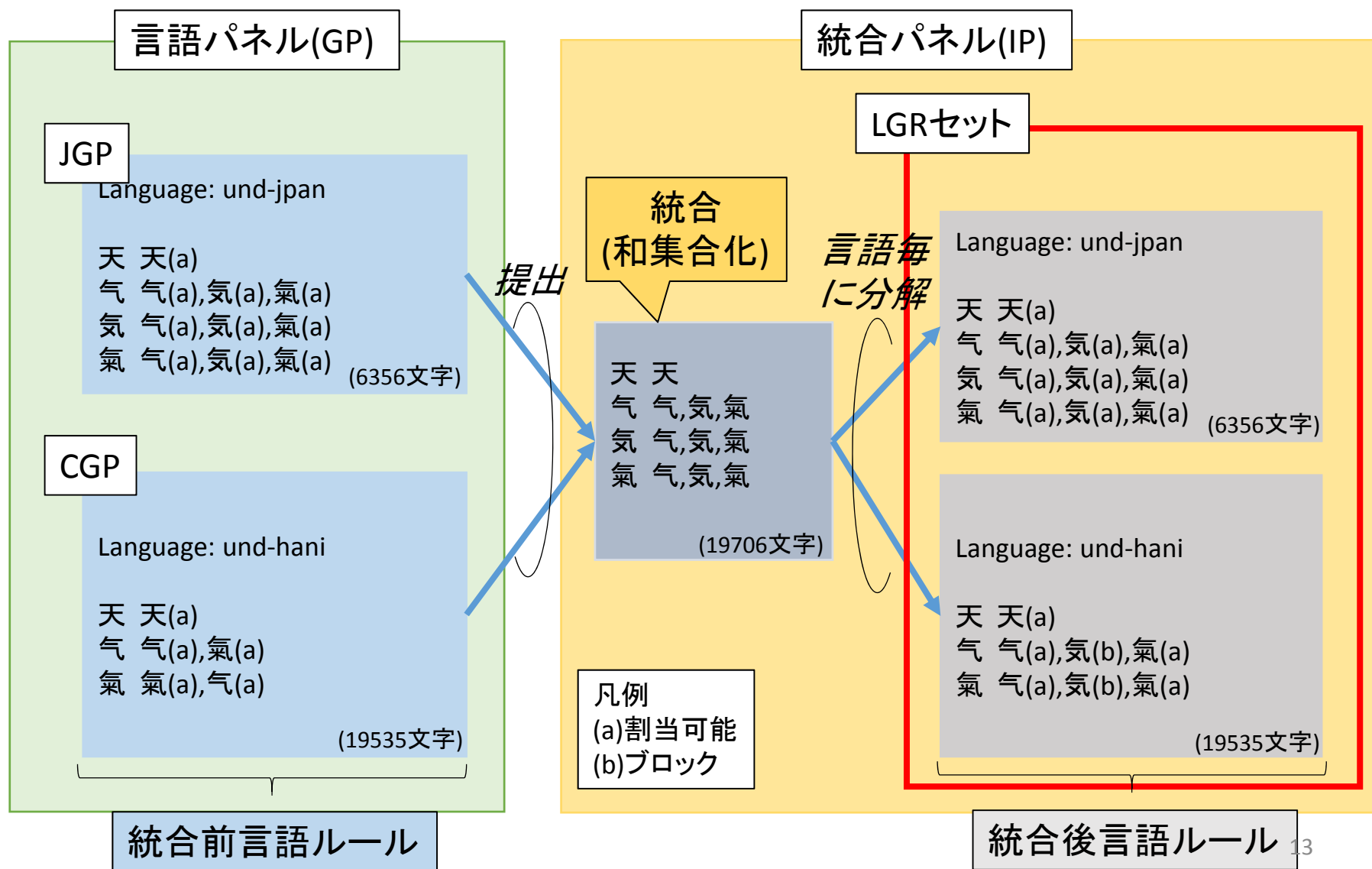
Language: und-jpan
申請文字列: 天氣
割当可能: 天氣
ブロック: 天气

Language: und-jpan
申請文字列: 天气
割当可能: 天气
ブロック: 天氣

Language: und-hani
申請文字列: 天气
割当可能: 天气, 天氣
ブロック:

Language: und-hani
申請文字列: 天氣
割当可能: 天气, 天氣
ブロック:

例4: JGPは積極的に異体字を定義する



例4: 文字列生成例

【注】

混在可能: (a)と(a)

混在不可能: (a)と(b)、(b)と(b)

Language: und-jpan
申請文字列: 天氣
割当可能: 天氣, 天氣, 天气
ブロック:

Language: und-jpan
申請文字列: 天氣
割当可能: 天氣, 天氣, 天气
ブロック:

Language: und-jpan
申請文字列: 天气
割当可能: 天气, 天氣, 天氣
ブロック:

Language: und-hani
申請文字列: 天气
割当可能: 天气, 天氣
ブロック: 天氣

Language: und-hani
申請文字列: 天氣
割当可能: 天气, 天氣
ブロック: 天氣