

(A) Měakaběnkadjy

En comparant des formes qui partagent des éléments, il est possible de classer les composantes de chaque mot de la façon suivante :¹

1	2	3	4	
ar	a	těm	mã	vous (peu nombreux) partirez
	i	nõr	ri	pendant que je me couche
ar		nõr		ils/elles (peu nombreux) se couchent
		těm	kadjy	il/elle est sur le point de partir
mě	i	bôx	mã	nous (plusieurs, sans toi) arriverons
	a	nõr	mã	tu te coucheras
mě		kaběn	ri	pendant qu'ils/elles (plusieurs) parlent
	ba	bôx	kadjy	nous deux (toi et moi) sommes sur le point d'arriver

À partir de cela, on peut voir les équivalences suivantes :

Pour la colonne (3) :

bôx	arriver
těm	partir
kaběn	parler
nõr	se coucher

Pour la colonne (1) :

(zéro)	singulier (sauf avec ba)
ar	peu nombreux
mě	plusieurs

Pour la colonne (4) :

(zéro)	présent
mã	futur
kadjy	être sur le point de
ri	pendant que

Pour la colonne (2) :

(zéro)	il/elle
i	je ou nous (sans toi)
a	tu ou vous
ba	nous (avec toi)

Avec ces informations, il est possible de traduire les expressions du problème :

- | | | |
|----|-----------------------|--|
| a. | aribôx | nous (peu nombreux, sans toi) arrivons |
| b. | akaběnmã | tu parleras |
| c. | abôxkadjy | tu es sur le point d'arriver |
| d. | kaběn | il/elle parle |
| e. | měbakaběnkadjy | nous (plusieurs, avec toi) sommes sur le point de parler |

¹On a placé les éléments dans des colonnes différentes en fonction de la possibilité qu'ils se combinent les uns avec les autres. Il est possible de faire autrement.

f.	ils/elles (plusieurs) sont sur le point de se coucher	mĕnōrkadjy
g.	je me coucherai	inōrmā
h.	il/elle part	tĕm
i.	pendant que nous deux (toi et moi) partons	batĕmri
j.	il/elle se couche	nōr

Les données de ce problème proviennent du travail de terrain d'Andrés Pablo Salanova sur le mĕbĕngōkre; les traductions ont été légèrement uniformisées pour ce problème. Il y a aussi une petite simplification dans **mĕnōrkadjy**, puisque ce verbe normalement change de forme quand le sujet est pluriel : il passe de **nōr** à **ikwā!** Il n'y a pas assez d'information dans le problème pour déterminer cela.

(B) Quatre-vingt-dix-neuf, cent huit

Il y a plusieurs chiffres qui ont la même structure :

pemp ha daou-ugent
tri ha tri-ugent
daouzek ha tri-ugent
dek ha pevar-ugent
naontek ha pevar-ugent
unan ha pevar-ugent

 (a) ha (b) -ugent

Notez que (a) et (b) semblent être tirés du même ensemble de chiffres, car **tri** apparaît dans les deux fonctions. Notez aussi qu'il y a des expressions complexes parmi les mots qui apparaissent dans (a): **daouzek** pourrait être formé de **daou** et **dek**, et ce dernier élément pourrait être présent aussi dans **naontek**.

Il est raisonnable de supposer que ces chiffres expriment des quantités dans la forme suivante : $(a) + (b) \times F$, où F est un facteur à déterminer (**ugent**). Il y a d'autres possibilités —par exemple, (a) pourrait exprimer des dizaines, et (b) des unités—, mais je ne peux penser à rien de plus simple que la première hypothèse. Si on se trompe, on finira par le savoir.

Il reste à déterminer la valeur de F . Pour cela, il est utile de se rappeler des valeurs dont il est question :

9 18 27 36 45 54 63 72 81 90 99 108

Sauf pour 90 et 99, il y a seulement une valeur par dizaine, ce qui rend impossible que F soit 10 (car il y aurait trois chiffres dans la dizaine de **pevar**). Par contre, si F est 20, les chiffres 81, 90 et 99 tombent tous les trois dans la quatrième vingtaine. Il n'y a pas de F plus petit qui permettrait cela.

On déduit alors que **pevar** est 4 (car **ugent** est 20 et **pevar-ugent** est 80). L'observation que **naontek** est composé avec **dek** nous permet de supposer que **dek** est 10 (alors **dek ha pevar-ugent** est 90) et **naontek** est 19 (**naontek ha pevar-ugent** est 99); maintenant **unan** ne peut être que 1 (**unan ha pevar-ugent** est 81).

Puisque nous avons 4 et 9, essayons de les trouver dans les autres chiffres : ils sont dans **pevar ha hanter kant**, et peut-être dans **nav**. Il est très raisonnable de supposer que ce dernier est 9, car il n'y a rien qui ressemble plus à **naou**. D'autre part, **pevar ha hanter kant**, s'il doit contenir un 4, ne peut être que 54. Cela voudrait dire que 50 échapperait à la logique de compter par vingtaines. Hmm. Alors il restent deux vingtaines qui contiennent deux quantités : 20 contient 27 et 36, 60 contient 63 et 72. Puisque **tri ha tri-ugent** est $x + x \times 20 = 21x$, il ne peut être que 63; **daouzek ha tri-ugent** est alors 72. D'où on déduit que **tri** est 3, **daou** est 2, et **daouzek** est 12. Puis, **pemp ha daou-ugent** est dans la vingtaine de 40, où il ne restait que 45; **pemp** est donc 5.

Maintenant, l'expression qu'on devrait traiter est **seizh warn ugent**, car on sait qu'elle est dans la vingtaine de 20. Puisqu'il n'y a pas de **dek**, **zek** ou **tek** dans la première partie, on peut supposer que **seizh** est moins de 10, alors 7. La présence de **warn** à la place de **ha** n'a pas d'explication. Mais parfois les chiffres plus basses ont des conventions légèrement différentes des autres.

L'expression **c'hwec'h ha tregont** devrait attirer notre attention maintenant. Elle a une structure semblable aux premières expressions que nous avons vues, mais à la place de (b)-**ugent** nous avons **tregont**, qui a une ressemblance avec **tri** et avec **ugent**. Peut-être alors que 30, comme 50, ne suit pas la convention de compter par vingtaines; **c'hwec'h ha tregont** serait alors 36, et **c'hwec'h** 6. Le mot **triwec'h**, qui a une structure tout à fait unique mais qui semble composé de **tri** et d'un raccourci de **c'hwec'h**, est vraisemblablement 18.

Nous avons trouvé des équivalences pour tous les chiffres sauf 108.

Les deux expressions qui restent, **kant eizh** et **kant nemet unan**, contiennent le mot **kant**. On reconnaît **unan**, mais on a épuisé les valeurs qui ont 1 comme unité. Par contre, il est possible que cela exprime $100 - 1 = 99$; **kant eizh** serait alors 108, ce qui fait que **kant** soit 100, et **eizh** 8. On comprend maintenant que **hanter kant** est la moitié de cent.

Voici donc les équivalences :

1	unan	10	dek, -tek, -zek
2	daou	12	daouzek
3	tri	18	tric'hwec'h
4	pevar	19	naontek
5	pemp	20	ugent
6	c'hwec'h, -wec'h	30	tregont
7	seizh	50	hanter kant
8	eizh	100	kant
9	nav, naon-		

Bien entendu, les ressemblances avec le français et autres langues indo-européennes vous ont sans doute aidés (**nav** = 9, **dek** = 10, **tri** = 3, **unan** = 1, **daou** = 2, **kant** = 100, **ugent** = 20, et peut-être d'autres)!

La structure normale des chiffres est : (**kant**) (a) **ha** (b)-**ugent**, avec les exceptions suivantes : **ha** devient **warn** si (b) est vide, et **tregont** et **hanter kant** sont employés pour les chiffres entre 30 et 39 et 50 et 59, respectivement. On peut supposer que **kant nemet unan** est un cas spécial, qui ne s'applique que à 99.

Cependant, j'avoue que je n'étais pas sûr d'avoir bien résolu ce problème avant de traduire les expressions suivantes :

- a. **pemp ha pevar-ugent + pevarzek = kant nemet unan**
85 + 14 = 99
- b. **triwec'h ha tri-ugent + unan ha hanter kant = kant nav warn ugent**
78 + 51 = 129

Suivant la logique décrite ci-dessus :

- c. 10 **dek**
- d. 22 **daou warn ugent**
- e. 38 **eizh ha tregont**
- f. 74 **pevarzek ha tri-ugent**
- g. 156 **kant c'hwec'h ha hanter kant**

Ce problème a été conçu par Ksénia Guiliarova, avec ses propres données.

(C) Le monde à l'envers

On observe que tous les mots en *gasó* contiennent les lettres **gas** (d'où le nom). En plus, la voyelle qui suit le **gas** répète celle qui précède (ou à l'envers), donc on peut considérer qu'une des deux est insérée. Essayons de voir si enlever ces lettres nous aide dans quelque chose :

- | | | | |
|--------------------|--------------------|---------------------|------------------|
| a. kami•sa | chemise | k. batido•r | délateur |
| b. baka•ñ | bon-vivant | l. pape•l | papier |
| c. fri•o | froid | m. afwe•ra | dehors |
| d. ade•ntro | à l'intérieur | n. aba•xo | en bas |
| e. taru•go | maladroit | ñ. kaba•co | cheval |
| f. korpi•ño | soutien-gorge | o. mari•do | mari |
| g. ami•go | copain | p. xi•l | personne crédule |
| h. tfori•so | saucisse | q. kabe•sa | tête |
| i. rebe•s | envers | r. ta•ngo | tango |
| j. peti•so | petit (en stature) | s. bomba•tja | culotte de femme |

Il devrait être possible de voir deux choses : (1) le • tombe après la dernière voyelle si celle-ci est suivie d'une consonne; autrement, il tombe après l'avant-dernière voyelle (cela n'est important que pour la

dernière partie du problème); (2) les mots en *vesre* contiennent presque les mêmes lettres que ce qu'on voit dans la liste précédente, mais en ordre différent. Si vous connaissez l'espagnol, vous verrez aussi que plusieurs de ces mots sont des mots espagnols courants, et que l'emplacement du • correspond à l'accent, mais cela n'ajoute rien à la résolution du problème.

On peut aligner ces mots avec les équivalents en *vesre* assez facilement :

1.	c. → 7	fri•o	ofri	3.	f. → 12	korpi•no	nokorpi
	l. → 20	pape•l	pelpa		h. → 18	tfori•so	sotfori
	r. → 6	ta•ngo	gotan		s. → 17	bomba•tfa	tfabomba
2.	a. → 15	kami•sa	samika	4.	j. → 2	peti•so	tisope
	e. → 5	taru•go	goruta	5.	m. → 3	afwe•ra	arafwe
	g. → 13	ami•go	gomia		n. → 9	aba•xo	axoba
	k. → 10	batido•r	dortiba	6.	p. → 19	xi•l	loxi
	ñ. → 14	kaba•co	cobaka	7.	b. → 4	baka•n	kamba
	o. → 16	mari•do	dorima		d. → 11	ade•ntro	atroden
	q. → 1	kabe•sa	sabeka		i. → 8	rebe•s	behre

Dans le premier ensemble, on voit tout simplement une inversion dans l'ordre des deux syllabes du mot.² Quand un mot contient plus de deux syllabes, le changement d'ordre peut suivre plusieurs patrons. Mais attention : le problème ne vous demande pas de savoir quand chaque patron est choisi. Juste pour nous amuser un peu, examinons cela de près.

Si on identifie les trois syllabes d'un mot trisyllabique comme ABC, le premier ensemble de mots trisyllabiques (2.) est dans l'ordre CBA en *vesre*, le deuxième (3.) dans l'ordre CAB, le troisième (4.) dans l'ordre BCA, et le quatrième (5.) dans l'ordre ACB. Il ne manque que BAC pour avoir tous les ordres possibles, ce qui suggère qu'en fait il n'y a pas de contraintes sur l'ordre. Moi même je ne sais pas ce qui fait qu'une inversion est préférée à une autre, mais j'imagine qu'on cherche à optimiser le rythme au sens large et des associations avec d'autres mots. Ce qui importe, pour le problème, ce sont les contraintes par rapport à ce qu'on peut faire.

Regardons maintenant les deux derniers ensembles. L'avant-dernier ensemble (d'un seul membre) nous montre que l'inversion doit cibler *des syllabes*, puisqu'un mot qui n'a qu'une seule syllabe doit subir l'ajout d'une voyelle (et donc devenir disyllabique) pour subir l'inversion. Le dernier ensemble exemplifie certains changements qui affectent les sons de l'espagnol selon leur position dans le mot : **m** et **n** deviennent **ŋ** à la fin d'un mot, tandis que **s** devient **h** devant une consonne.

Maintenant on peut passer à la partie plus difficile du problème :

²Il n'est pas nécessaire de savoir ce qu'est une syllabe ni l'endroit où l'espagnol sépare un mot en syllabes pour résoudre ce problème, mais il est utile d'y penser. Quand un mot consiste en une alternance de consonnes et voyelles (CVCV), chaque CV constitue une syllabe. Quand il y a plus d'une consonne de suite, les règles sont plus complexes. Les données du problème montrent que la séparation de syllabes dans **tango** et **afwera**, par exemple, est **tan.go** et **a.fwe.ra**, car leur inversion est **gotan** et **arafwe** et non pas **ngota** et **afrawe**.

t.	pjegasesa	→	pje•sa	→	sapje	
u.	gigasita	→	gi•ta	→	tagi	
v.	tigasinto	→	ti•nto	→	totin	→ totiŋ
w.	pregaseso	→	pre•so	→	sopre	
x.	pagasantfo	→	pa•ntfo	→	tʃopan	→ tʃopaŋ
y.	pagasansa	→	pa•nsa	→	sapan	→ sapaŋ

Chacun des changements représentés par une flèche a été expliqué ci-dessus.

Finalement, pour trouver des mots en *gasó* à partir du *vesre*, il faudra tenir compte de l'emplacement variable du • :

z.	mjoŋka	→	kamioŋ	→	kamio•ŋ	→	kamiogasoŋ
aa.	xabru	→	bruxa	→	bru•xa	→	brugasuxa

Les données du *vesre* pour ce problème sont tirées de mes propres connaissances, en consultation avec Agustín Anselmi. Pour le *gasó*, les principes sont tirés de l'entrée *Rosarigasino* du Wikipedia en espagnol.

(D) Test d'articulation

Si on examine l'ensemble de données complètes, on peut assez facilement établir les correspondances suivantes entre la prononciation adulte et la prononciation de l'enfant en question :

[p]	→	b, p	bɔm, xup	[s]	→	x	bɛ̃xɛx
[b]	→	b	bɔ̃	[z]	→	ʧ	ʧɛb
[m]	→	m	mubɛt, bym	[ʃ]	→	x	goxɔ̃
[f]	→	b	bɔk	[ʒ]	→	ʧ	nay
[v]	→	b, ʧ	bax, ʧɛɣ	[l]	→	j, w	ejebã, xawad
[w]	→	w	xwax	[j]	→	j	gɔɣij
[t]	→	d, t	dɛt	[k]	→	g, k	gœ, bɔk
[d]	→	d	done	[g]	→	g	gɔɣij
[n]	→	n	done	[ʁ]	→	ʧ	ʧɛɣ

Il y a aussi quelques cas où une consonne se perd dans la prononciation de l'enfant. J'ai choisi de montrer les correspondances des séquences de consonnes où cela se produit plutôt que d'ajouter un ∅ (ce qui veut dire zéro) dans les correspondances ci-dessus, ce qui aurait rendu leur description plus compliquée. On voit que des consonnes ne disparaissent dans la prononciation de l'enfant que quand elles font partie d'une séquence de consonnes dans la prononciation de l'adulte!

[pl]	→	b	bɛ̃
[bl]	→	b	bœ
[pʁ]	→	b	bɛ̃xɛx
[bʁ]	→	b	ʧɛb
[fl]	→	b	bœɣ

[vr] → **y** **jiy**
 [vw] → **b** **bayyy**
 [tʁ] → **d** **dē**
 [dz] → **y** **yiγ**
 [ʃj] → **x** **xē**
 [kl] → **g** **gɔx**

Il nous reste alors à déterminer les contextes pour les différentes correspondances lorsqu'il en existe deux chez l'enfant, par exemple p → **b** (**bɔm**), **p** (**xup**). Pour p, t, k et v, il semble assez clair que la première variante se trouve devant des voyelles, tandis que la deuxième se trouve à la fin du mot. Pour l la chose semble un peu plus compliquée, mais on peut dire qu'il devient j sauf devant un a. On espère que les questions ne nous pousseront pas trop à rendre la règle plus précise.

Voyons maintenant les correspondances :

a.	petit	[pətsi]	bəxi
b.	grand-père	[gʁɑ̃pɛʁ]	gãbɛγ
c.	treize	[tʁɛz]	dɛγ
d.	chèvre	[ʃɛvʁ]	xɛγ
e.	frère	[fʁɛʁ]	bɛγ
f.	lampe	[lɑ̃p]	wāp
g.	cafard	[kafaʁ]	gabay
h.	fini	[fini]	bini
i.	château	[ʃato]	xado
j.	j'aime	[ʒem]	γem
k.	toutes	[tut]	dut
l.	frisée	[fʁize]	biγe
m.	saucisse	[sosis]	xoxix
n.	cadre	[kadʁ]	gad (*)
o.	poisson	[pwasɔ̃]	baxɔ̃
p.	lit	[li]	ji

(*) Il n'y a pas de correspondant pour [dʁ] dans nos données, alors la réponse à (n) ne s'offre pas automatiquement! Qu'est-ce qu'on fait ? Ici, je choisis d'assimiler ce cas au cas qui lui ressemble le plus, celui de [bʁ] dans *zèbre*, qui produit **b** à la fin du mot.

Ce problème a été conçu par Harold Somers, avec ses propres données.

(E) Ne te fige pas devant ce problème

La première partie de ce problème est un exemple d'un type très commun de tâche. Elle se résout en comptant le nombre de répétitions d'un certain mot dans les deux parties des données. Ici, la présence de deux marqueurs, **e** et **a**, le premier lié au verbe (on le sait car il n'y en a qu'un seul par phrase),

le deuxième lié aux substantifs (car il peut y en avoir un ou deux par phrase), nous aide à trouver la structure de base.

- | | |
|------------------------------------|---------------------------------|
| a. e vacu a gone | 8. l'enfant frappe |
| b. e unuma a tii a marama | 5. la femme boit du thé |
| c. e la'o a gone | 6. l'enfant va |
| d. e voro'a a yame a tagane | 10. l'homme casse la lame |
| e. e vacu'a a marama a gone | 4. l'enfant frappe la femme |
| f. e lo'i a yame | 3. la lame est tordue |
| g. e corita a waqa a marama | 9. la femme attache le canot |
| h. e unu a gone | 1. l'enfant boit |
| i. e luaca a waqa a gone | 2. l'enfant vomit dans le canot |
| j. e voro a waqa | 7. le canot est cassé |

On voit aussi que certains verbes apparaissent sous deux formes, selon qu'ils ont ou pas un complément; entre parenthèses, j'ajoute des formes qui apparaissent plus tard dans le problème, mais qui suivent le même patron :

	<i>sans complément</i>	<i>avec complément</i>
frapper	vacu	vacu'a
boire	unu	unuma
aller	la'o	(la'ova)
casser	voro	voro'a
tordre	lo'i	(lo'ia)
vomir		luaca
attacher	(cori)	corita

Avec cela, on peut résoudre la deuxième partie :

- | | |
|------------------------------------|--------------------------|
| k. e la'o a marama | la femme va |
| l. e voro'a a waqa a marama | la femme casse le canot |
| m. e cori a waqa | le canot est attaché |
| n. e la'ova a yame a marama | la femme va la lame (??) |
| o. e lo'ia a yame a tagane | l'homme tord la lame |

On voit bien qu'il est facile de trouver la forme sans complément à partir de celle qui a un complément, mais non pas dans l'autre sens... il est impossible de déterminer si la terminaison qui sera ajoutée sera **'a, ma, va, ca, ta, ou a**. On a six possibilités différentes pour sept paires de formes! Mais on sait que si on a une forme du verbe avec complément, la forme du verbe sans complément se trouve en retirant la dernière voyelle et consonne (s'il y en a).

Cela rend la résolution de la troisième partie relativement facile :

