

# HINDI

Annie Montaut

INALCO, EHESS-CNRS

## 1. Contexte historique

Alors que jusqu'au VII<sup>e</sup> s., le hindi médiéval se caractérise par la spécificité sectaire de ses parlers (braj lié à Krishna et son poète Surdas, awadhi à Ram et son poète Tulsidas, *sant bhāsā* à la divinité abstraite), le hindi standard moderne représente, plus que la continuation d'un parler régional littéraire spécifique, le développement, chez les lettrés de la cour, puis la standardisation par sanscritisation, d'une *lingua franca*, ou d'une *koinè* (l'hindustani) dérivée de ces parlers et parlée dans la région de Delhi-Agra.

## 2. Écriture

Le système graphique du hindi, qui est à quelques simplifications et additions près celui du sanscrit (cf. sanscrit et nagari), n'est qu'abusivement considéré comme phonologique, plus abusivement encore comme syllabique. Certes, l'ordre alphabétique est déterminé par les lieux et modes d'articulation : voyelles, puis occlusives (des vélares aux dentales en cinq séries respectives : sourde, sourde aspirée, sonore, sonore aspirée, nasale), enfin fricatives, liquides et semi-consonnes ensemble. Mais la correspondance graphe-son n'est pas toujours bi-univoque : certains symboles correspondent à une combinaison de plusieurs sons (pour *kṣ*, *tr*, *gy*, le *jña* du sanscrit) et certaines distinctions graphiques sont neutralisées dans la prononciation moderne (les deux chuintantes *ṣ* et *ś* se prononcent [ʃ], l'ancien *r̥* voyelle comme la séquence *ri*). Surtout, la voyelle *a* réalisée en sanscrit après toute consonne non abrégée et non suivie de voyelle s'élide en hindi à peu près comme le *e* muet français. Une consonne finale n'est suivie

du *a* que précédée d'au moins une autre consonne. La deuxième consonne comme la première dans *samajh*, écrit *s + m + jh*, est syllabique, mais non dans *samjhā*, écrit *s + m + jh + ā*.

Outre les deux symboles *ṛ* et *ṛh* notés par un point souscrit sous les graphes représentant les dentales sonores rétroflexes, le hindi a aussi complété les signes simples de la *nāgarī* pour y accommoder les sons de l'arabe et du persan, passés à divers degrés dans la langue : *f*, *z*, *q* (dorsale uvulaire), *ǰ* (*y* : fricative dorsale voisée non aspirée), *x* (fricative dorsale sourde non aspirée) sont ainsi notés par un point souscrit aux symboles graphiant *p<sup>h</sup>*, *j*, *k*, *g*, *k<sup>h</sup>*. Cela donne un total de quarante-quatre graphes de consonnes simples non modifiés, trois combinaisons (sans entrée spécifique dans le dictionnaire) et sept signes modifiés par un point souscrit, dont deux indigènes (soit cinquante-quatre signes).

Voyelles et consonnes ont deux formes, indépendante et dépendante. La voyelle qui suit une consonne est notée par un signe diacritique placé après (*ā*, *ī*, *o*, *o*), avant (*i*), au-dessous (*u*, *ū*) ou au-dessus (*e*, *ε*) de cette consonne. La nasalisation est notée par un point suscrit sur un croissant (*candra-bindu*) ou par un simple point suscrit (*anusvāra*) si le signe de la voyelle dépasse la ligne sous laquelle s'écrivent les lettres. Le graphe d'une consonne enchaînée à la suivante (sans le *a* « inhérent ») s'abrège de façons diverses mais régulières. Cependant, ce traitement graphique ne rend compte que d'une faible partie des cas où la prononciation réalise des successions CC, souvent fonction de règles phonotactiques et morphologiques sans graphie particulière. Mais la norme orthographique est absolument contrainte (*akasmāt* s'écrit avec le symbole abrégé du *s*, *asmān* avec le symbole plein). À cette différence phono-graphique majeure avec le sanscrit, s'ajoute l'extension de l'usage de l'*anusvāra*, pour noter en particulier les consonnes nasales précédant immédiatement une consonne de même articulation. Quant à la ponctuation moderne, elle ajoute au *daṇḍa*,

barre verticale traditionnelle fonctionnant comme point, les signes occidentaux.

### 3. Phonétique, phonologie

L'inventaire des consonnes met en évidence dix-sept occlusives, huit fricatives, quatre affriquées, huit nasales, six liquides (quarante-trois consonnes), susceptibles de se géminer, et trois semi-consonnes. La transcription usuelle est indiquée après le symbole phonétique.

	Uvulaire	Pharyng.	Vélaire	Palatale	Rétroflexe	Alvéolaire	Dentale	Labio-dentale	Bilabiale
<i>Occlusive</i>									
<i>sourde</i>	q		k		t ṭ		t		p
<i>sourde asp.</i>			k <sup>h</sup>		t <sup>h</sup> ṭ <sup>h</sup>		t <sup>h</sup>		p <sup>h</sup>
<i>sonore</i>			g		ɖ ɗ		d		b
<i>sonore asp.</i>			g <sup>h</sup>		ɖ <sup>h</sup> ɗ <sup>h</sup>		d <sup>h</sup>		b <sup>h</sup>
<i>Affriquée</i>									
<i>sourde</i>				tʃ c					
<i>sourde asp.</i>				tʃ <sup>h</sup> c <sup>h</sup>					
<i>sonore</i>				dʒ j					
<i>sonore asp.</i>				dʒ <sup>h</sup> j <sup>h</sup>					
<i>Fricative</i>									
<i>sourde</i>			x	ʃ ṣ		s		f	
<i>sonore</i>		h	ɣ ġ			z			
<i>Nasale</i>									
<i>non asp.</i>			ŋ ñ	ɲ ñ̃	ɳ ɳ̣		n	m	ɱ
<i>aspirée</i>							n <sup>h</sup>	m <sup>h</sup>	
<i>Liquide</i>									
<i>latérale</i>					l ḷ	l			
<i>latérale asp.</i>						l <sup>h</sup>			
<i>trill</i>						r			
<i>flap</i>					ɽ ɽ̣				
<i>flap aspiré</i>					ɽ <sup>h</sup> ɽ̣ <sup>h</sup>				
<i>Semi-cons.</i>				y				v	w

L'inventaire des voyelles en met en évidence vingt-cinq : treize orales et douze nasales.

Aperture		Avant		Arrière			
		orale	nasale	orale non arrondie	nasale non arrondie	arrondie orale	arrondie nasale
-	brève	i	ĩ			u	ũ
	longue	ī	ĩ̄			ū	ũ̄
	brève	e	ẽ			o	õ
	longue	ē	ẽ̄	ə ə̄		ō	ȭ
+	brève			a	ã		
	longue	ɛ ai	ẽ̄ aĩ̄	ā	ã̄	ɔ au	ȭ aũ̄

Les traits distinctifs des voyelles sont la hauteur (+ haut et + bas, les voyelles moyennes étant - haut - bas), la postériorité (+/-), la longueur (*ki* « que », *kī* « de », *sūt* « coton », *sut* « fils », *kām* « travail », *kam* « moins »), la tension (toute voyelle tendue est longue), la nasalisation.

Parmi les traits distinctifs des obstruantes, les plus remarquables sont : la rétroflexion (*dāl* « lentille », *ḍāl* « branche », *sāth* « avec », *sāṭh* « 60 »), l'aspiration (*sāt* « 7 », *sāth* « avec » ; *mātā* « mère », *māthā* « front » ; *pal* « instant », *phal* « fruit » ; *sabā* « vent d'est », *sabhā* « assemblée », *pāk* « pur », *pākh* « quinzaine claire »), l'affrication (*śor* [ʃor] « bruit », *cor* [tʃor] « voleur »). Plusieurs arguments suggèrent de traiter ce trait comme distinctif, et l'affriquée comme un seul son (tʃ, dʒ / c, j) plutôt que comme une séquence dentale palatale : l'analyse en deux phonèmes contraindrait à introduire de nouvelles occlusives palato-alvéolaires, et trois nouvelles fricatives (ʃ<sup>h</sup>, ʒ, ʒ<sup>h</sup>, / ś<sup>h</sup>, ž, ž<sup>h</sup>), solution peu économique et peu conforme à l'économie du système. Les consonnes affriquées et aspirées sont mieux analysées comme un son et non deux (CC) car elles se comportent comme une consonne simple dans les alternances morphophonologiques qui entraînent la non-réalisation du *a* dit inhérent : de même qu'on a *aurat* « femme », *aurtẽ* femmes, *saṛak* rue, *saṛkẽ* « rues », *nikal* sors, *niklā* « sorti »,

on a *pighal* « fond », *pighlā* « fondu », alors que la séquence C<sub>1</sub>C<sub>2</sub> dans le même contexte n'entraîne pas l'amuïssement du *a* avant C<sub>3</sub> (*pustak* « livre », *pustakē* « livres », *nartak* « danseur », *nartakō* « danseurs », *kucalnā* « écraser », *kuclā* « écrasé »).

Ces obstruantes contrastent dans toutes les positions, sauf les occlusives *ḍ* et *ḍh* qui n'apparaissent pas en position intervocalique et ne contrastent pas en finale dans les mots indigènes, et les deux chuintantes *ś* et *ṣ*, qui se prononcent identiquement. Parmi les sons empruntés (à l'arabe, via le persan : *ġ*, *x*, *q*, *z*, *f*), seuls les deux derniers sont pleinement intégrés dans le système phonologique, correspondant à une opposition sonore/sourde régulièrement réalisée par ailleurs (*falak* « œil », *palak* « paupière », *phalak* lame ; *sajā* « orné », *sazā* « punition » ; *rāj* « royaume », *rāz* « secret ») ; les autres s'opposent dans la prononciation des connaisseurs de l'ourdou (*bāg* « reine », *bāġ* « jardin », *xam* « tournant », *ġam* « tourment », *ham* « nous », *xānā* « lieu », *khānā* « manger », *qasāi*, « boucher », *kasāi* « attache »), et l'arrêt glottal n'est réalisé qu'exceptionnellement, même par les ourdouphones (*'ām* « commun », *ām* « mangue », *kām* « travail »). Quand ils ne sont pas réalisés, c'est le phonème indigène le plus proche qui les remplace (*kasāi* « boucher », *bāg* « jardin », *khuś* pour *xuś* « heureux »).

Parmi les sonantes, *m*, *n*, *r*, *l*, *y*, *v* contrastent dans toutes les positions, mais non le flap *r* et les aspirées *n<sup>h</sup>*, *l<sup>h</sup>*, *r<sup>h</sup>* (pas à l'initiale), *m<sup>h</sup>* (ni en finale ni à l'initiale) : *kumār* « prince », *kumhār* « potier ». La nasale rétroflexe ne contraste que dans de rares paires (*bān* « habitude », *bāṇ* « flèche », *bām* « terrasse »). Elle représente le plus souvent un des six allophones de *n* conditionné par la nature de la consonne suivante (ici rétroflexe : *ṭhaṇḍ* « froid », *ghaṇṭā* « heure »). De même les nasales vélaire *ṅ*, palatale *ñ* et labiodentale *ṁ* n'apparaissent que devant consonne homorganique (*sañjh* « soir », *pañjī* « liste », *vañś* « lignée », *śaṅkh* « conque », *saṅg* « ensemble », *saṅgh* « association », *saṅvlā* « ocre »). Dans de tels contextes d'assimilation, *n* et *m* fonctionnent aussi comme des allophones dental et

bilabial (*lambā* « long », *gandh* « odeur », *gandā* « sale », *panth* « voie », *hans* « oie »), alors que, dans les autres positions, elles s'opposent clairement (*mātā* « mère », *nātā* « lien », *kām* « travail », *kān* « oreille », *bān* « habitude », *bām* « terrasse »), ce qui n'est pas le cas des trois autres nasales. C'est aussi le cas de *s* qui s'assimile en *ṣ* devant rétroflexe (*kaṣṭ* « problème », mais *durust* « correct, sain », *goṣṭ* « viande ») comme *l* (*uḷṭā* « renversé », *kheltā* « jouant »).

Les flaps *ɽ* et *ɽh* sont aussi en quasi-distribution complémentaire avec les occlusives rétroflexes *ɖ* et *ɖh* : sauf dans les emprunts, les flaps interviennent entre voyelles, avant et après une consonne non rétroflexe (*pīṛā* « souffrance », *uṛṭā* « s'envolant », *lakṛī* « bois »), et en finale après voyelle (*bāṛ* « enclos », *bāṛh* « inondation »), les occlusives rétroflexes à l'initiale (*ḍor* « corde », *ḍhor* « bétail ») et avant rétroflexe ou semi-consonne (*dhanāḍhya* « riche »).

Autres allophones : *w* (arrondi) précédé d'une consonne dans la même syllabe, *v* (non arrondi mais non consonantique) autrement (*swayā* « soi », *nahatwatā* « importance », *dwārā* « par », *itvār* « dimanche », *vah* « il », *avinās* « indestructibilité ») après *ā* (*rāv*, *rās* « roi »).

Les voyelles fermées *e* et *o* sont brèves devant la séquence -hC, ou h final, longues ailleurs : *mōhinī* « séduisante », *mohlat* « loisir ». Dans le même contexte, le *a* inhérent se réalise *e* (*pahlā* [pehlā] « premier », *rah* [reh] « reste », *rahā* « resta ») : *e* est allophone de *a* devant -hC ; -ō final est long. L'avancement du *a* jusqu'au *e* central (ɐ, ə) correspond également à des allophones de *a* (*aglā* « prochain », mais *bas* [bʌs] « assez », *par* [pʌr] « sur », *lahar* [ləhar] « vague »). Le *a* final (ə surbref) après CC ne se réalise que dans les emprunts sanscrits après CC (*svapna*).

Au total, vingt-cinq obstruantes (sans compter l'arrêt glottal ni *ǰ*, *x* et *q*, réalisés en ourdou, ni *ʃ*), douze sonantes (en comptant *ɳ*, *ɽ* et *ɽh*), et vingt voyelles sont phonologiques en hindi standard.

#### 4. Structure syllabique

La tendance, clairement manifestée dans les cas de resyllabification, est à la syllabation ouverte maximale (*pāp* « péché », *pā.pī* « pécheur ») avec une attaque maximale et une coda minimale, la contrainte de sonorité progressive (*sonority sequence*) rendant compte de la plupart des configurations CCVCC : comme la syllabe hindi est structurée par une montée progressive de la sonorité de l'attaque au pic sonore (la voyelle qui est le noyau de la syllabe) et une descente progressive à partir du pic, la première consonne est normalement moins sonore que la deuxième, l'avant-dernière plus que la dernière, sur une échelle allant des occlusives (1) aux sonnantes, aux semi-voyelles (9) et à la voyelle (10). Ainsi les groupes initiaux *glāni* « remords », *bhram* « illusion » et *tyāg* sont-ils permis et non *\*lga*, *\*rbh*, *\*yta*. Ils sont aussi contraints par le principe d'accroissement sonore minimal (plus d'un degré de différence est requis entre C1 et C2), et l'évitement des séquences de consonnes articulées au même point (*r* faisant exception). La coda, elle, n'obéit qu'au premier principe, acceptant entre autres des consonnes proches dans l'échelle de sonorisation, des géminées et des consonnes homorganiques (*gupta* « caché », *sikkh*, *yuddh*, « guerre », *σᾶνγ* « compagnie »). À l'attaque comme dans la coda, les sifflantes échappent à ces contraintes (*snān* « bain », *stabdh* « sidéré », *strī* « femme »), mais les séquences sC(C) à l'attaque sont généralement resyllabifiées dans le parler non standard à l'aide d'une voyelle prothétique : *is.trī* « femme », *as.nān* « bain », *is.kūl* « école ».

L'accent est fonction du poids de la syllabe. Une syllabe est lourde à partir de deux mores (voyelle longue ou voyelle brève suivie de consonne), mais les syllabes de deux mores, voire de trois, sont traitées comme légères si elles voisinent dans le mot avec une syllabe de quatre mores (V longue CC par ex.). L'accent tombe sur la syllabe la plus lourde du mot (*ka'lī* « bourgeon », *ban.'dūk.bā.zī* « armé »). Si les syllabes sont de poids égal, il tombe sur la pénultième (*'dhva.ni* « son », *'bar.tan* « plat », *sā.'lā.nā*

« annuel »), et s'il y a plusieurs syllabes lourdes, sur l'avant-dernière lourde (*ba.'hā.du.rī* « bravoure », *ma.hā.'bhā.rat*). Les mots fléchis ont un comportement accentuel de mots simples (*'la.ghu* « bref », *la.ghu.'tar* « plus bref », *'ghar* « maison », *gha.'rō* « maisons ») comme les dérivés (*'sun.dar* « beau », *sun.'dar.tā* « beauté »), les redoublements onomatopéiques et les mots « écho » (*'śā.dī-vā.dī* « mariage et tout »), mais non les redoublements complets ou sémantiques, où chaque mot garde son accent (*'pī.lā-'pī.lā* « jaunâtre » ou « bien jaune », *'kām-'kāj* « occupation »). L'accent est fixe et prédictible.

## 5. Morphologie

### 5.1. Catégorie nominale : cas, genre et nombre

Le hindi, langue à deux genres et deux nombres, analytique à fortes tendances agglutinantes, n'a que de très faibles vestiges flexionnels. Il distingue toujours la forme oblique du nom (devant postposition ou en fonction de circonstant sans postposition) au pluriel par un suffixe -ō. La flexion ne subsiste, pour distinguer forme directe (sujet, objet direct) et oblique, ou singulier et pluriel, que pour les noms masculins en -ā (qui fléchissent dans ces deux cas en -e). Les féminins n'ont pas d'oblique singulier et forment leur pluriel par un même suffixe -ē (-ā̃ après -i-). Les différences de traitement sont conditionnées par la phonétique du mot, les masculins indigènes en -ā (ā̃) seuls ayant une véritable flexion, les autres (consonne, ī, u, ū, o, mots sanscrits en -ā comme *rājā* « roi », *pitā* « père ») ne varient que par la suffixation à la base de -ō, comme les deux types de féminins, ceux en ī (-i, -u) et ceux en consonne (ou mots sanscrits en -ā comme *bhāṣā* « langue »). Bien que ces marques soient d'origine flexionnelle, elles ne se présentent plus en synchronie que comme des suffixes, avec d'éventuels sandhis (ī s'abrègeant devant voyelle avec *yod* épenthétique : *lar̥kī* sg., *lar̥kiyā̃*, *lar̥kiyō* pl. ; ū dans le même contexte s'abrège). Seul le premier type de masculin présente une alternance



flexionnelle au sens étroit, notée sans [-] dans le tableau, à la différence des suffixes.

	Masculin en ā		Autres masculins		Féminin en ī, -i, -u		Autres féminins	
	sg.	pl.	sg.		sg.	pl.	sg.	pl.
Forme directe	ā	e	ø	ø	ī (-i, u)	-ã	ø	-ẽ
Forme oblique	e	õ	ø	-õ	ī (-i, u)	-õ	ø	-õ

L'adjectif varie moins que le nom, seul le premier type (en ā) se fléchit (en e), les autres adjectifs étant invariables. Il marque ses degrés syntaxiquement ou lexicalement (*sab se* « tous » abl. + adj., *bohut* « très ») sauf dans des survivances (*priyatam* « très cher » < SCR, ou *badatar* « pire » < PERS).

Les variations du nom ne se substituent pas aux cas du sanscrit ; c'est la postposition, adjointe au nom et non à ses déterminants, qui indique la fonction de tout le groupe nominal : *ne* ergatif, *ko* datif et accusatif différentiel (humain ou objets spécifiques), *mẽ* locatif, *se* ablatif/instrumental. Le génitif est spécial : la postposition qui sert à le former, à l'origine suffixe adjectival, prend la marque de genre, nombre et cas du nom recteur (*kā, kī, ke* : « de »). Ainsi : *acchī laṛkī* « fille bonne », *acchī laṛkiyõ kā ghar* « la maison des filles bonnes », *acchā ghar* « bonne maison » (masc.), *acche ghar* « bonnes maisons », *acche gharõ mẽ* « dans les bonnes maisons », *laṛkī ke gharõ mẽ* « dans la maison de la fille » (mais *sundar/khālī gharõ ko*, comme *sundar/khālī ghar*, « belle/vide maison »).

Les locutions postpositionnelles sont le plus souvent issues d'adverbes auxquels s'adjoint la marque de « génitif » : (*ke*) *bād* « après », (*ke*) *sāmne* « en face (de) », *ke madhyam* « au milieu (de) », (*ke*) *ūpar* « au-dessus (de) », (*ke*) *nīce* « sous », (*ke*) *pīche* « derrière », etc.

Le genre et le nombre ne se distinguent souvent que par les accords (adjectif et participe), la distinction -ā/ī ne valant que pour une classe limitée

de noms, les suffixes féminins formant aussi une classe limitée (*-nī, -in, -ānī*, etc.). Mais l'importance des marques d'accord dans le paradigme verbal interdit de considérer les catégories du genre et du nombre comme faibles en hindi standard.

### 5.2. Les pronoms

Dans le paradigme des pronoms, la catégorie de la personne et du nombre a partiellement cédé à celle de la hiérarchie de respect : il y a trois formes pour la 2<sup>e</sup> personne (familiale *tū*, informelle *tum* et honorifique *āp*) ainsi que pour la 3<sup>e</sup> (*vah*, familier et informel, *ve*, honorifique, *āp*, sur-honorifique), *tum*, *āp* et *ve* valant pour les deux nombres. La forme *vah* et son pluriel *ve* fonctionnent aussi comme démonstratif distal. Le genre n'est pas marqué, ni dans le démonstratif marquant la proximité (*yah*, pl. *ye*) ni dans le relatif (*jo*). Les pronoms ont des formes obliques flexionnelles (*maī/mujh*, *tū/tujh*, *vah/us*, *ve/un*, *yah/is*, pl. *ye/in*, *jo/jis*, pl. *jo/jin*). Seuls *ham*, *tum* et *āp* ne varient pas. Le réfléchi (*apne āp*, *āp*), qui renvoie au sujet de la proposition, est invariable. Les postpositions s'amalgament aux pronoms, avec diverses variantes stylistiques : *mujhko* ou *mujhe*, *tujhko/tujhe*, *usko* ou *use* (et le non standard *mere ko*, *tere ko*). On peut considérer que les bases en *u-* ou *v-*, en *y-* ou *i-* des démonstratifs sont aussi des allomorphes de la même unité, réalisée par la semi-consonne ou la voyelle selon la nature du son suivant. Les possessifs sont la forme (adjectivale) du génitif du pronom (*uskā*, *āpkā*, *jinkā*) sauf pour les vrais personnels qui ont des formes en *-r-* (*merā*, *terā*, *tumhārā*, *hamārā*) et le réfléchi (*apnā*).

### 5.3. Le verbe

Le paradigme verbal a pour base la forme de l'impératif singulier. L'infinitif en est dérivé comme un nom masculin, qui peut fléchir (*māke kah-ne ke anusār* [mère / de / dire-inf. obl. / d'après] « selon le dire de Mère », de *kah-nā* « dire »). Les deux participes accompli et inaccompli,

adjectivaux, suffixent la désinence de genre et de nombre directement au radical (*kah-ā* « dit ») ou après la marque *-t-* d'inaccompli (*kah-t-ā* « disant »). Un troisième participe invariable sert à coordonner deux procès (*kah-kar*). La base connaît une flexion radicale, qui oppose intransitif à transitif, ou transitif à causatif, par la tension de la voyelle (*ruk* « s'arrêter », *rok* « arrêter » ; *piṭ* « être battu », *pīṭ* « battre » ; *nikal* « sortir », *nikāl* « expulser »), mais aussi par des modifications consonantiques (*tūṭ* « se casser », *toṭ* « casser »), ou l'adjonction d'un suffixe *-ā* avec la tension de la voyelle radicale (*bīt* « se passer », *bitā* « passer » ; *sīkh* « apprendre », *sikhā* « enseigner » ; *so* « dormir », *sulā* « endormir »). C'est toujours un suffixe (*-vā*) qui forme le factitif (faire faire : *rukvā* « faire arrêter », *sulvā* « faire endormir », *sikhvā* « faire enseigner »), procédé convergent sinon emprunté au dravidien qui a trois formes de bases verbales correspondant à trois degrés de causation.

Le paradigme de conjugaison est entièrement régulier et analytique, structuré par une opposition aspectuelle de base opposant participe inaccompli (*-t-*) et accompli (*-ø-*). Les temps de l'indicatif sont tous composés sauf l'aoriste (dont la forme est analogue au participe accompli, sans valeur temporelle), le participe portant l'aspect, et l'auxiliaire (« être » : *honā*), toujours postposé, portant le temps : *āī* [venir-f.sg. (= ppe acc.)] « elle vint », *vah paṛhtā hai* [3sg. / étudier-ppé inacc. m.sg. / être-prés.-3sg.] « il étudie », *vah paṛhtā thā* [3sg. / étudier-ppé inacc. m.sg. / être-passé-m.sg.] « il étudiait ». À l'inaccompli, on distingue obligatoirement l'habituel (ppe inacc. aux.-temps) et le progressif (base rester-ppé acc. aux.-temps) : *maī us din skūl jā rahā thā* [1sg. / ce-obl. / jour / école / aller / progr.-m.sg. / être-m.sg.] « ce jour-là, j'allais à l'école ». Le futur est formé par la suffixation de *-gā* < « aller » sur l'ancien présent, aujourd'hui subjonctif, seul tiroir verbal qui présente, avec les désinences personnelles (*-ū̃*, *-o*, *-e*), des vestiges de flexion, toutefois neutralisés au pluriel (*-ē* aux trois personnes). La distinction +/- accompli s'étend à tous les temps composés

(subjonctif passé et irréel passé) formés sur les participes, les formes modales simples étant respectivement issues de l'ancien présent synthétique pour le subjonctif (*kahũ* « que je dise ») et analogue au ppe inacc. pour le contrefactuel, sans valeur temporelle : *mai yahã baithtĩ* [je / ici / s'asseoir-irr. (= ppe inacc.) -f.sg.] « je m'assiérais, me serais assise ».

De nombreux auxiliaires marquent optionnellement diverses nuances aspectuelles (ou d'*aktionsart*). Ils qualifient le déroulement du procès, comme le duratif (ppe inacc. + rester : *vah paṛhtā rahtā hai* [3sg. / étudier-ppe inacc. m.sg. / durat.-ppe inacc. m.sg. / être-3sg.] : « il étudie sans répit, n'arrête pas d'étudier », le duratif-progressif (ppe inacc. + aller / venir / marcher), l'habituel du progressif (*honā* « être » sous sa forme longue), ou les marges du procès (inchoatif : inf. + toucher ; terminatif : base + être soldé). Fréquentatif (ppe acc. + faire) et permissif (inf. + donner) sont grammaticalisés. « Pouvoir » a deux expressions (*pānā* « parvenir », *saknā* « être capable »), « devoir », trois (*cāhie*, *honā*, *paṛnā*, selon que l'obligation est générale, ponctuelle et future, objet d'une contrainte externe). L'opposition +/- perfectif, d'acquisition récente, est en cours de grammaticalisation : une douzaine de verbes de mouvement désémantisés à un certain degré, dits « explicateurs », s'adjoignent à la base nue (aller, tomber, se lever, donner, jeter, prendre), et transforment un imperfectif en perfectif, en orientant aussi la diathèse (*tū apnā pāṭh paṛh le* [réfl.-m.sg. / leçon-m.sg. / étudier / prendre-imperf.] « étudie ta leçon »).

L'indicatif est nié par la particule *nahĩ*, les autres modes par *na* (l'impératif par *na* ou *mat*).

La personne n'est donc marquée qu'au présent (-ũ 1sg. et -o 2 seules sont distinctives), les autres formes ne s'accordant qu'en genre et en nombre, sur les participes.

Le passif (auxiliaire *jānā* « aller » postposé au participe accompli) touche aussi bien les intransitifs (avec, en contexte négatif, la valeur d'incapacité) que les transitifs : *mujhse yahã baithā nahĩ jātā* [1sg.-instr. / ici / s'asseoir-

acc. / nég. / passif-m.sg. / « je ne peux pas m’asseoir là » (c’est plus fort que moi).

#### 5.4. Dérivation

La dérivation est surtout féconde dans la création néologique technique (vocabulaire scientifique, politique, administratif, économique) et fait exclusivement appel à des affixes sanscrits sur des bases sanscrites : aux préfixes et suffixes persans (*be-*, *-i*, *-dar* : *be.imān* « malhonnête », *imān.dār* « honnête », *imān.dār-ī* « honnêteté », *be.xabar.ī* « insouciance », *ġarīb.ī* « pauvreté ») ou arabes (*lā.davā* « sans remède », *ġair.qānūn* « illégal »), qui ne sont plus du tout productifs, on substitue les très nombreux affixes sanscrits : *ni-* et ses allomorphes phonétiquement conditionnés *niṣ-*, *nís-*, *nir-*, *nis-* (la forme *nís-* devant sourde palatale, *niṣ* devant sourde vélaire ou labiale, *nis-* devant sourde sifflante, *nir* ailleurs – voyelle, sonore) dérive, comme *a-*, un adjectif privatif d’un nom, *-tā* dérive un nom abstrait, *-ak* un nom d’agent : *nís.cint.tā* « insouciance », *nir.mal* « sans tache, pur », *a.sthir.tā* « instabilité », *darś.ak* « spectateur »). Certains formations affixées obéissent aux règles du sanscrit (allongement de la première voyelle radicale avec le suffixe adjectivant *-ik* (*din* « jour », *dainik* « quotidien », *samāj* « société », *sāmājik* « social »).

Dans la langue courante, le suffixe *-vālā* est le seul vraiment productif, fournissant toutes sortes de relations et notamment la possession, l’origine, l’agent : *hindīvālā* « hindiphone », *ġāvvālā* « villageois », *lakṛīvālā* « en bois », *dekhnevālā* « spectateur » (suffixé à un adjectif, il le spécifie à l’intérieur d’un ensemble : *pīlīvālī* « la jaune », *lambāvālā* « le grand »). La marque de féminin sert de diminutif : *rassā* « corde épaisse, câble », *rassī* « petite corde », *jholā* « sac, besace », *jholī* « petit sac, pochette ».

### 5.5. Composition

Les diverses formes de redoublement et de paires synonymiques relèvent du lexique. Le redoublement du lexème entier, réputé expressif quand il n'est pas distributif (*garam-garam* « tout chaud, bien chaud »), joue le rôle d'un intensif mais avec les adjectifs de couleur, d'un atténuatif (*nīlā-nīlā* « bleuâtre ») comparable au suffixe *-sā* (< *jaisā* < *sadriśya* comme : *nīlā-sā* « bleuâtre »).

Les verbes dits composés ou locutions verbo-nominales, classe plus importante que celle des verbes simples, relèvent aussi de la composition tout en jouant le rôle d'un procédé de dérivation très productif pour fournir un prédicat à partir d'un adjectif ou d'un nom (*garam honā/karnā* « chauffer » intr./trans., *bāt karnā* [parole faire] « parler » ; *bhūk honā/lagnā* [faim être/toucher] « avoir faim »). Un verbe étranger s'emprunte comme formant nominal d'une locution verbale (*join karnā* « rejoindre », *caik karnā* « vérifier »). Ces locutions, dont le formant nominal est plus ou moins autonome au niveau des accords, permettent, par le choix du formant verbal (verbe support ?), de diversifier très souplesment la voix, l'aspect, la diathèse : *yād* « souvenir », combiné avec *honā* « être », est statique, avec *rahnā* « rester », duratif, avec *ānā* « venir », inchoatif, avec *karnā* « faire » actif, avec *rakhnā* « poser » actif duratif (« garder en mémoire »).

## 6. Syntaxe

Comme dans la plupart des langues indiennes, le prédicat est final dans la phrase, les compléments les plus lointains (circonstants) étant les plus éloignés du verbe, les plus proches (objet) le précédant immédiatement. L'élément recteur suit l'élément régi (dét. adj. N, N2 N1, V aux., etc.), le comparé suit le marqueur de comparatif qui suit le comparant : *Sītā kusum se lambī hai* [Sita Kusum abl. grande est] « Sita est plus grande que Kusum ». Mais la marque d'interrogation totale, comme les subordinants, est en position initiale, et la négation préverbale : *kyā vah mere dost ke ghar nahī*

*āegī* [interr. / 3sg. / mon / ami / de / maison / nég. /venir-fut.-3f.sg.] « ne viendra-t-elle pas chez mon ami ? »

Autre trait panindien : un complément d'objet humain ou spécifique est marqué par la postposition *ko* (dat./accus.) : *larkī ko bulāo* [fille / *ko* / appelle] « appelle la fille ».

La phrase simple n'est jamais nominale (sauf quand elle est exclamative) : la copule (*honā*) a la même forme que le verbe existentiel formant les prédications d'existence et de localisation (il y a X, X est dans la maison). La structure de type nominatif (SV) ou nominatif-accusatif (SOV), qui oppose les verbes intransitifs (dormir, marcher) et transitifs (frapper, lire) n'est pas dominante. Trois autres cas se présentent dans l'énoncé élémentaire.

L'agent du verbe transitif à l'accompli est oblique et marqué spécifiquement (*ne*) comme ergatif, le verbe s'accordant avec le patient direct et prenant, si celui-ci est marqué, la désinence « par défaut » du masculin singulier : *hamne yah pustak likhī* [1pl.-erg. / ce / livre-f.sg. / écrire-f.sg.] « nous avons écrit ce livre ». Locutions verbo-nominales : c'est l'abstrait verbal (nom d'action) qui se comporte comme « patient » (plus justement, thème) dans les locutions faiblement coalescentes (et liées au participant externe par une postposition « de » : *kā intazār karnā* « faire attente de ». Autrement, les deux formants, s'ils sont coalescents, fonctionnent comme un verbe simple qui s'accorde avec le participant externe (*pasand karnā*). Si l'objet est marqué, l'accord est neutralisé (m.sg.) : *hamne is pustak ko dekhā hai* [nous-erg. / ce / livre / dat. / voir / est] « nous avons vu ce livre (en question) ».

Quel que soit le nombre des participants, un prédicat « expérientiel », c'est-à-dire renvoyant à des états physiologiques, psychiques ou cognitifs, requiert le datif de l'expérient et s'accorde avec l'autre terme non marqué : *mujhe bhūkh lagī hai* [1sg.-dat. / faim-f.sg. est collée-3f.sg.] « j'ai faim »,

*baccō ko hindī ātī hai* [enfant-m.pl. / dat. / hindi-f.sg. / venir-f.sg.] « les enfants connaissent le hindi ».

Le possesseur, en l'absence d'équivalent lexical pour « avoir », est au locatif (dans, près de) ou au génitif selon que la possession est acquise ou non : *uske do bhāī haī* [3sg.-gén. / deux / frères / sont] « il a deux frères », *uske pās tīn kamīzē thī* [3sg.-de / près / trois / chemises-f.pl. / étaient-f.pl.] « il avait trois chemises », *unmē kāfī sāhas thā* [3pl.-loc. / assez / courage-m.sg. / était-m.sg.] « ils avaient beaucoup de courage ».

Les deux derniers modèles (typiques de l'expérience subjective et de la possession) s'étendent à de très nombreuses locutions verbo-nominales, le participant au datif impliquant une absence de contrôle volontaire, voire de conscience réflexive, le participant au génitif correspondant plus à la stativité. Les déontiques ont ainsi un participant principal au datif, outre les verbes subjectifs, et les expressions suivantes au génitif : *sādī honā* « épouser », *irādā honā* « avoir l'intention », *sampark honā* « avoir contact » mais aussi *bāt honā* « parler », *bahas honā* « discuter » : *merī unse bāt hūī* [1sg.-gén. / 3pl.-abl. / parole-f.sg. / fut-f.sg.] « je leur ai parlé »).

De tels types d'énoncé mettent en évidence la faible pertinence des catégories grammaticales de sujet et d'objet. Les traits sémantiques (contrôle volontaire, stativité) mais aussi pragmatiques (on euphémise le participant au génitif en le périphérisant : *unkā dehānt huā* [3pl.-gén. / mort / fut.]) priment sur l'organisation syntaxique. Cependant, dans chacun de ces cas, c'est le participant principal, le premier dans la chaîne séquentielle, qui contrôle la réflexivation et la coordination même s'il n'est pas exprimé : *(mujhe) apnī bahan kī khabar milkar sidhā calī gāī* [1sg.-dat. / réfl. / sœur / de / nouvelle / se-trouvant / (je) droit / partis] « apprenant ces nouvelles de ma sœur, je partis sur le champ » ; *yah sunkar baṛī khuśī huī* [ceci / entendant / grand / bonheur / fus] « je fus ravi d'entendre ceci » ; *apne ko dekhkar (mujhe) dar laga* [réfl. / acc. / voyant (1sg.-dat.) / peur / toucha] « je pris peur en me voyant ». Malgré la saillance sémantico-pragmatique du premier



participant, c'est le terme le moins saillant dans une relation asymétrique (le patient par rapport à l'agent, le thème par rapport à l'expérient, le possédé par rapport au possesseur) qui est traité comme pôle de l'énoncé (non marqué et régissant l'accord verbal).

La coordination est exprimée par des conjonctions (*aur* « et », *lekin* / *par* / *magar* / *kintu* « mais ») ou par l'usage du participe conjonctif pour le premier prédicat : *vah andar ākar baiṭh gāi* [3sg. / dedans / venant / s'asseoir / alla-f.sg.] « elle entra et s'assit ».

Pour marquer les relations de dépendance entre propositions, le mode privilégié est l'emploi d'un mode non fini pour le verbe dépendant, les participes redoublés pouvant exprimer diverses relations avec le verbe principal (concomitance, cause-effet, simultanéité, évitement), l'infinitif suivi de diverses postpositions équivalant à diverses circonstancielles (*X ke jāne se pahle* « avant que X ne vienne », *ke kāraṇ* « parce que X est venu »). Le système corrélatif (*jo... vah* « lequel... celui-là », « qui », *jab... tab*, « quand, alors », *jahā... vahā* « où, là », *jaise... vaise*, « comme, de même ») est si prégnant que les conjonctions empruntées au persan sont souvent anaphorisées ou cataphorisées par un déictique ou un adverbe (*agar... to* « si, donc », *hālāki... phir bhī, cūki... islie* « puisque, c'est pourquoi », *yah... ki*, « ceci... que »). C'est aussi la corrélation qui exprime le comparatif d'égalité : *merā darzī jitnā amīr hai, māi utnā hī hū* [mon / tailleur / autant / riche / est / je / autant / juste / suis] « je suis tout aussi riche que mon tailleur ».

## 7. Lexique

La modernisation par recours au sanscrit a abouti, en hindi standard, à une énorme proportion de mots *tatsam*, « restés tels », parfois en concurrence avec les *tadbhav*, « devenus tels ». Ainsi *Kṛṣṇa karm* « travail », *grām* « village », *pustak* « livre », *varṣā* « pluie », coexistent avec *Kānh*, *kām*, *gāv*, *pothā*, *barsāt*, qui en sont respectivement issus, le registre stylistique,

plus ou moins formel, justifiant le choix. Mais certains champs sémantiques n'ont qu'un membre, le *tadbhav* dans le domaine rural (*ũṭ* « chameau »), le *tatsam* dans le domaine technique (*śvetāṇu* « leucocyte », *dvinetrī* « jumelles »), administratif (*nivedan patr* « requête ») et métalinguistique (*dantoṣṭhya* « labiodentale »). Certains des *tadbhav*, encore courants au XIX<sup>e</sup> s. (Kellogg, p. 41), ont reculé devant la sanscritisation massive du hindi standard et sont aujourd'hui dialectaux ou archaïques (ainsi *lāh* « profit », *koh* « colère », remplacés par les *tatsam lābh* et *krodh*).

Les emprunts étrangers sont massivement issus du persan et de l'arabe (via le persan) en hindi, quoique aujourd'hui taris, alors qu'en ourdou, la langue technique se forge à partir de l'arabe (d'où la divergence des deux langues, à l'origine proches au point de se confondre dans le parler courant : Nehru, de culture linguistique hindoustani, ne comprenait ni la version hindi ni la version ourdou de la Constitution indienne). Les termes arabes ou persans sont souvent plus courants que leur équivalent sanscrit : ainsi *ādmī* (AR) vs *puruṣ* (SCR) « homme », *xatm* (AR) vs *samāpt* (SCR) « fini », *kitāb* (AR) vs *pustak* (SCR) « livre », *lekin* (AR) vs *kintu* (SCR) « mais », *zamīn* (PERS) vs *bhūmi* « terre », *zindā* (PERS) vs *jīvit* (SCR) « vivant », *pareśānī* (PERS) vs *cintā* (SCR) « ennui ». Certains emprunts au persan se différencient peu de l'indo-aryen (*stān* « lieu » < IR vs *sthān* < SCR) et les deux bases peuvent coexister dans des sens un peu différents (*band* « fermé » (PERS) et *bandh* (SCR) « attacher »), comme du reste les *tatsam* et les *tadbhav* (*megh* « nuage », *meh* « pluie »).

Une centaine de mots sont d'origine turque (premières invasions entre le VIII<sup>e</sup> et le XIII<sup>e</sup> s). De la colonisation occidentale restent quelques termes portugais (*cābī* « clef », *gobhī* « chou-fleur ») mais c'est surtout l'anglais qui continue à enrichir la langue.

Outre le calque et l'emprunt, un des procédés d'enrichissement du lexique, ancien et panindien, a consisté à forger des mots sur des formations onomatopéiques, directement imitatives ou « expressives », redoublées ou

non, suffixées souvent par *-ak* : *kal-kal* « bruit d'un cours d'eau qui coule », *ṭap-ṭap* « bruit d'un écoulement goutte à goutte », *ṭapaknā* « gouter, dégringoler », *khaṭkhaṭā* « frapper (à la porte) », *gun-gun* « bourdonnement », *gungunānā* « fredonner », *jham-jham* « briller » (verbe *jhamaknā*), *cam-cam* « scintillement » (verbe *camaknā*), *ḍabḍab* « sensation des larmes qui montent aux yeux » (verbe *ḍabḍabānā*), *dharaknā* « palpiter », *chūn-chūn* « tintement des bracelets », etc. Dans les mots dits « à écho », autre structure panindienne, un *v-* se substituant à la première consonne du lexème redoublé (*cāy-vāy* « thé », etc., *corī-vorī* « vol », etc.), la formation désigne un ensemble qui englobe et déborde le champ sémantique de la notion de base (« et ce qui va avec »). La composition, par ailleurs, associe souvent deux synonymes ou complémentaires (*ghūmnā-phirnā* [se promener-tourner] « se balader », *mā-bāp* [mère-père] « parents », *pati-patnī* [mari-femme] « couple ») et, comme le redoublement (*kabhī* « une fois », *kabhī-kabhī* « parfois »), est un procédé traditionnel d'enrichissement lexical.

## Bibliographie

- Kellogg Rev. H. S., *A Grammar of Hindi Language*, Delhi, Oriental Book Reprints, 1872 (rééd. : 1975).
- Montaut A., *Hindi Grammar*, Munich, LINCOLM Europa (Indo-European Linguistics Series), 2004.
- *Voix, aspects et diathèses*, Peeters, 1991.
- Six articles sur le hindi in Montaut (dir.), *Les Langues d'Asie du Sud* (Ergativité : p. 57-64, Temps, aspect, mode : p. 209-246, Dérivation causative : p. 215-220, Locutions verbo-nominales : p. 221-228, Constructions expérientielles et possessives : p. 229-238, Particules, p. 239-248, Bibliographie), Ophrys, 1997.
- Ohala M., *Aspects of Hindi Phonology*, Delhi, Motilal Banarsidass, 1983.
- Singh R. et Agnihotri R. K., *Hindi Morphology. A Word-Based Description*, Delhi, Motilal Banarsidass, 1997.