

# Indicateur tarifaire de l'IA et directives contractuelles DACH

Publié par l'Association des speakerines et speakers professionnels de Suisse VPS-ASP

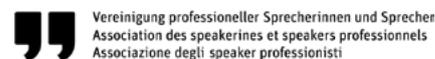
Cet indicateur est né de la collaboration des associations germanophones de speakerines et speakers.



VDS Allemagne



VOICE Autriche



VPS-ASP Suisse

sur la base des directives globales AI Guidelines for Voice Work



UVA – United Voice Artists  
[www.unitedvoiceartists.com](http://www.unitedvoiceartists.com)

En complément de la [liste des tarifs de la VPS-ASP](#), l'indicateur tarifaire de l'IA permet de s'orienter dans le calcul des tarifs des speakerines et speakers dans le nouveau segment de la synthèse vocale (voix IA).

## Contenu

---

Préambule

1. Principes contractuels généraux
2. Synthèse de la voix
  - 2.1 Frais d'apprentissage neuronal
  - 2.2 Rémunération de la synthèse de base
3. Rémunération lors de l'utilisation de l'IA
4. Sonorisation minimale en direct
6. Autres aspects contractuels
  - 6.1 Aspects juridiques
  - 6.2 Aspects contractuels



## Präambel

---

### **“L’utilisation de l’IA doit élargir l’épanouissement humain et non le réduire. L’IA ne doit pas remplacer l’homme“**

Alena Buyx, présidente du Conseil d’éthique allemand (2023)

La communauté internationale des speakerines et speakers est convaincue que les œuvres basées sur les émotions humaines et la communication entre les personnes doivent être interprétées par de véritables voix humaines. Nous pensons que seuls les êtres humains sont capables de traduire les émotions en mots et de créer un lien empathique avec le public ou celles et ceux qui les écoutent. L’intelligence artificielle (IA) ne peut ni maîtriser ni même comprendre les nuances nécessaires à une exécution de cette tâche en termes de qualité et d’authenticité.

Nos associations sont également conscientes du fait que le développement des voix synthétiques en combinaison avec l’IA progresse inexorablement et qu’il est donc nécessaire de l’encadrer et de l’évaluer correctement afin de fournir à tous les acteurs et actrices du marché une orientation et une base de calcul cohérentes, ainsi qu’une protection contre les abus.

Les possibilités techniques de production et d’utilisation de l’IA et des voix de synthèse se développent à un rythme de plus en plus rapide - et les droits des artistes sont souvent négligés. En Europe notamment, le RGPD (Règlement Général sur la Protection des Données, ndt), les droits de la personnalité, les droits d’auteur, les droits à rémunération, la responsabilité, etc. protègent les artistes et leur travail, tandis que dans d’autres pays, la protection juridique varie considérablement.

La technologie vocale de l’IA risque, dans le pire des cas, de détruire la base économique des professionnels de la voix et de les priver du contrôle de l’exploitation de leur propre voix. C’est pourquoi les associations conseillent vivement à toutes les speakerines et tous les speakers du monde entier d’examiner soigneusement et consciencieusement les contrats qu’on leur propose, en particulier en ce qui concerne l’utilisation et l’application de l’IA.

Les directives et recommandations suivantes ont pour but d’aider toutes les parties concernées dans le monde entier (par exemple les artistes, les agents, les commanditaires, les donneurs et preneurs de licence, les agents de marque, les avocats et les législateurs) à établir des accords contractuels et des rémunérations appropriées pour les projets linguistiques utilisant l’IA. Ce document a donc pour but de servir de guide pour des négociations équitables.



# 1. Principes contractuels généraux

---

Afin de garantir une collaboration équitable, les principes suivants devraient constituer la base de tout contrat.

## Consentement général

La création d'une empreinte numérique de la voix de toute personne ne peut se faire qu'avec son consentement explicite. Une description spécifique et détaillée de l'utilisation prévue doit être fournie.

## Rémunération adaptée

La rémunération doit refléter l'étendue de toutes **les utilisations prévues** de la copie numérique de la voix d'un/e speaker/ine, en tenant compte de tous les aspects (quantitatifs, temporels, spatiaux, techniques, médiatiques, etc).

## Limitation générale de l'utilisation

Des **droits d'utilisation** totalement illimités pour l'utilisation de la copie numérique de la voix ne devraient jamais être consentis.

## Limitation dans le temps

Il ne faut jamais convenir de droits illimités dans le temps pour l'utilisation du duplicat vocal numérique.

## Principe d'opt-out

Le/La speaker/ine doit avoir la possibilité de mettre fin à l'utilisation de sa voix pour la création d'IA à la fin de chaque période contractuelle. Dans ce cas, le preneur de licence (donneur d'ordre) doit garantir, à la demande du donneur de licence (speaker/ine), que les données relatives à sa voix seront intégralement supprimées et, en outre, qu'il empêchera la production ultérieure de cette voix ou de ses caractéristiques par des moyens techniques appropriés.

## Fidélité vocale (Blended Voices / Morphing)

En raison de la traçabilité indispensable, de la cogestion, des droits de la personnalité et de la publicité, le « Blended Voices / Morphing » est fermement rejeté par les speakerines et speakers. Si un tel accord est malgré tout conclu, chaque speaker/ine individuel/le dont la voix fait partie de la voix hybride doit être indemnisé intégralement et non pas proportionnellement. Il doit également conserver un droit de consentement complet et être en mesure, en toutes circonstances, de contrôler le répertoire dans lequel la voix hybride est utilisée.

## Fidélité linguistique (unicité de la langue, langues étrangères)

Afin de protéger les marchés internationaux et de préserver la culture de chaque langue, les traductions de la langue parlée générées par l'IA ne devraient pas être effectuées.

## Principe du lieu de commercialisation

Afin d'éviter le dumping salarial dû à des niveaux de rémunération différents sur les marchés internationaux, les rémunérations devraient toujours être au moins égales aux normes du pays d'exploitation.



## 2. Synthèse de la voix

---

**Pour que les systèmes d'IA soient en mesure de générer des voix artificielles, la voix originale d'un/e speaker/ine doit d'abord être synthétisée.**

Chaque speaker/ine est bien entendu libre de synthétiser sa propre voix à ses frais et de la proposer en en assurant lui-même la gestion.

Cependant, si la synthèse de la voix est effectuée par un commanditaire / un studio / une production, le/la speaker/ine doit être rémunéré/e pour cette prestation. Ce chapitre présente les différents éléments de rémunération appropriés.

En outre, un contrat complet tenant compte des principes contractuels généraux (chapitre 1) concernant l'utilisation de ces enregistrements est absolument indispensable.

Les postes concernant la rémunération de ce chapitre 2 « Synthèse de la voix » ne concernent que la saisie de la voix dans le système d'IA et/ou le travail en studio, ou l'option consistant à proposer la voix. Ils n'accordent aucun droit d'utilisation. L'utilisation doit toujours faire l'objet d'une rémunération supplémentaire. (voir chapitre 3)

Le « Neural Learning Fee » (2.1) et la « rémunération pour la production de matériel de formation » (2.2) doivent être considérés comme des paiements uniques. Le « Listing Fee » (2.3), en revanche, doit être traité comme un paiement récurrent (par exemple mensuel).

### 2.1 Frais d'apprentissage neuronal - (Neural Learning Fee)

---

Chaque fois qu'une voix humaine avec son identité vocale spécifique est introduite dans le réseau neuronal d'un fournisseur d'IA (par ex. Open AI, Eleven Labs, Microsoft, Google), ce système bénéficie d'un apprentissage. Ce processus constitue la base du développement des systèmes d'IA neuronaux à commande vocale et sa valeur est élevée.

#### **Pour mieux comprendre :**

*Le grand problème de l'apprentissage neuronal des systèmes d'IA réside dans le fait qu'une fois que l'apprentissage a eu lieu, il ne peut plus être annulé. Le système connaît ensuite un développement durable, même si les données d'entraînement initiales sont supprimées. En ce sens, le matériel fourni par des speakers/ines professionnels/les est d'une grande valeur pour les systèmes de formation.*

*Comme ce processus d'apprentissage va au-delà de la simple imitation du son d'une voix, puisqu'il englobe toute l'identité vocale d'une personne, c'est-à-dire tous les paramètres individuels de son caractère (par exemple la tonalité, la vivacité, la respiration, la fougue, l'expérience de la vie, l'humour, le charme, la gentillesse, la personnalité...), le/la speaker/ine met à disposition toute son expertise professionnelle et son statut économique. Pour cette raison, l'apprentissage neuronal peut potentiellement entraîner une perte économique totale pour le/la speaker/ine.*

*Ce qui est problématique, c'est que de nombreux commanditaires ne savent pas comment et où les données de formation qu'ils introduisent dans leurs systèmes sont ensuite traitées.*

Si ce processus d'entraînement ne peut pas être exclu de manière claire et démontrable par le commanditaire, le consentement explicite doit être obtenu et le/la speaker/ine doit être rémunéré/e par un « neural learning fee », qu'il s'agisse de matériel préenregistré ou d'enregistrements réalisés pour la circonstance.

**Pour cette raison, les associations conseillent en général de refuser clairement l'apprentissage neuronal ou, s'il ne peut être exclu avec certitude, de lui donner une valeur élevée - de l'ordre de six chiffres.**



## 2.2 Rémunération de la synthèse de base

---

Si un commanditaire souhaite créer un double vocal numérique d'un/e speaker/ine, il doit en demander l'autorisation de principe et rémunérer ce/tte dernier/ère par un cachet de base unique. Le cachet de synthèse de base comprend un forfait de 5 jours de studio pendant lesquels la voix du comédien/de la comédienne peut être enregistrée pour la synthèse. Tout temps de studio supplémentaire doit être rémunéré en sus.

<b>Synthèse de base forfaitaire</b> (5 jours de studio inclus)	CHF	4'250 - 6'500
<b>Tarif journalier</b> pour la lecture de matériel d'entraînement supplémentaire	CHF	850 - 1'300

### Licences et remarques :

*Ces licences s'appliquent uniquement à la synthèse de la voix dans le système fermé du commanditaire, à l'exclusion explicite de l'apprentissage neuronal. Ce tarif n'inclut aucun autre droit ou licence, que ce soit pour l'utilisation de la voix numérique générée par l'IA, l'utilisation des résultats de cette même voix ou l'apprentissage neuronal. Ces droits doivent être achetés en plus, si nécessaire. Le cachet de base est toujours dû dans son intégralité, à titre de forfait, quel que soit le temps passé en studio : journée complète, partielle ou nulle. Une journée de studio est limitée à 5 heures de travail.*

## 2.3 Frais de référencement (Listing Fee)

---

### Offre et vente de voix synthétisées

L'évolution de la technologie vocale de l'IA offre à différents acteurs du marché la possibilité d'étendre leurs services en proposant des voix numérisées (par exemple, les studios d'enregistrement, les bases de données de voix, les agences, etc). L'offre de voix numériques doit être préalablement autorisée par les speakers/ines en personne et rémunérée par des « frais de référencement » - (« listing fee »).

Le montant de ces « frais de référencement » est librement négociable et peut aller d'un cachet symbolique de 150 francs par an à plusieurs milliers de francs. Ces frais autorisent exclusivement le fournisseur à proposer l'image de voix synthétisée. Les droits sur la voix reproduite numériquement et sa production restent cependant toujours la propriété du/de la speaker/ine d'origine. Les frais de référencement n'incluent pas les droits d'utilisation. Tous les droits d'utilisation doivent être négociés et faire l'objet d'une licence, indépendamment des frais de référencement et de la compensation éventuelle pour la voix synthétisée.

Les frais de référencement devraient en outre toujours être accompagnés d'un contrat, qui règle tous les détails de la collaboration, la rémunération des licences d'utilisation et des accords clairs sur la transparence, le contrôle de la qualité et le consentement autour de l'utilisation des voix.



### 3. Rémunération lors de l'utilisation de l'IA

---

Ce chapitre traite de l'utilisation des mandats générés par l'IA. En principe, la rémunération des speakers/ines est basée sur la vente de licences d'utilisation de leur voix. La technologie de l'IA ne change rien à ce principe ! Principe que voici :

#### UTILISER, C'EST UTILISER.

Le montant de la rémunération dépend toujours de l'ampleur de l'utilisation et du groupe cible atteint. Les associations de plusieurs pays publient à cet effet des listes nationales déjà établies qui, en plus des montants concrets des rémunérations, contiennent des recommandations pour les forfaits, les rabais de quantité et d'autres aspects des rémunérations :

VPS ASP Suisse	Liste des tarifs VPS-ASP	<a href="https://vps-asp.ch/about/tariff">vps-asp.ch/about/tariff</a>
VDS Allemagne	VDS-Gagenkompass	<a href="https://sprecherverband.de/service/vds-gagenkompass">sprecherverband.de/service/vds-gagenkompass</a>
VOICE Autriche	Taux d'honoraires VOICE	<a href="https://sprecherverband.at/de/honorarsaetze">sprecherverband.at/de/honorarsaetze</a>

#### Pour une meilleure compréhension :

*Même si l'application de l'IA peut éventuellement permettre au commanditaire d'économiser du temps et de l'argent dans le processus de production, les droits découlant de l'octroi de licences d'utilisation de sa voix restent inchangés, quelle que soit la méthode utilisée pour la produire. Tant que la qualité d'une réplique numérique répond aux exigences d'un commanditaire en matière de représentation de sa marque, de son produit ou de son contenu avec une voix générée par l'IA, il n'y a aucune raison de rémunérer le donneur de licence (speaker/ine) différemment qu'auparavant.*

Étant donné que la technologie de l'IA peut non seulement offrir toute la gamme des formes déjà connues, mais encore de nouveaux types et volumes d'exploitation de la voix, les éléments concrets des listes de rémunération nationales ne sont parfois pas directement applicables. Dans ce cas, les listes nationales et leurs niveaux de rémunération dans les genres concernés doivent servir de base pour déterminer une nouvelle rémunération, en tenant compte des « principes contractuels généraux » énoncés au chapitre 2.



## 4. Sonorisation minimale en direct

Lors de l'utilisation de voix d'IA, une base de matériel vocal humain original doit obligatoirement être produite pour chaque production.

Les associations germanophones ne considèrent pas la synthèse vocale par l'IA comme un substitut à la voix réelle, mais comme un complément à celle-ci. Pour cette raison, et afin de conserver une diversité de styles dans différentes productions et pour différentes marques, un nombre minimum de sujets / émissions / chapitres / modules devrait être sonorisé et rémunéré par la voix réelle dans la plupart des projets, et seulement ensuite complété et rémunéré par sa voix d'IA dans le même style et la même expression.

Genre	Umfang Mindestvertonungen live	Anmerkungen
Paid Media	5 - 15 Sujets	par campagne / produit par ex. spot TV, spot radio, spot en ligne, etc.
Unpaid Media	1 sujet / 1 module /2 Min.	par campagne / produit p.ex. vidéo web, film d'entreprise, présentation, app, etc.
Livre audio 60 Min	1 FAH (Final Audio Hour)	par livre audio
Jeux	1 heure de studio	par rôle et par titre Longueur du texte par heure : 600 mots.  Applicable uniquement aux rôles monotones et exigeants sur le plan vocal, et jusqu'à 10 pickups.  Pas d'utilisation de l'IA pour les enregistrements de rôles réguliers.
Doublage	1 heure de studio	par rôle et par titre  Applicable uniquement pour un maximum de 10 pickups.  Pas d'utilisation de l'IA pour les rôles réguliers.
Contenu éditorial	15 minutes par film, épisode, émission, titre	Par exemple, documentaires, reportages TV, etc.

Tous les genres non mentionnés ici ne nécessitent pas nécessairement un nombre minimum de parties parlées en direct et peuvent être générées exclusivement par la voix de l'IA, sous réserve d'accord. Les licences et les remarques pour chaque genre sont indiquées dans les listes de cachets nationales.

Si une campagne est constituée de plusieurs médias différents, le nombre minimum de parties parlées du segment de prix le plus élevé s'applique (par exemple, pour les médias payants et les médias non payants, le nombre de parties parlées des médias payants s'applique).



## 5. Autres aspects contractuels

---

Lors de la négociation d'un contrat d'IA, les facteurs suivants devraient être examinés et pris en compte de manière approfondie.

### 5.1 Aspects juridiques

---

#### Droits de la personnalité et de publication - Le droit à sa propre voix

Le/la speaker/ine est particulièrement protégé par le droit de la personnalité. Ce qui inclut le droit à sa propre voix. Ce droit n'est pas transmissible, même en cas de reproduction de la voix. Le/la speaker/ine doit donc avoir la possibilité de refuser toute utilisation des enregistrements qui irait au-delà de l'objectif initialement convenu, même si une rémunération supplémentaire est proposée.

#### Droit d'auteur et propriété intellectuelle

Il devrait être clairement communiqué et écrit que le/la speaker/ine qui « prête » sa voix détient également les droits sur sa copie numérique. Dans l'UE et en Suisse, les speakers/ines sont protégés/es par le droit d'auteur et les droits de propriété intellectuelle, et seules les licences d'utilisation peuvent être transférées. Toutes les parties concernées doivent également être conscientes que le contenu généré par l'IA n'est pas automatiquement protégé par les droits d'auteur.

#### Protection des données

Le titulaire de la licence doit s'assurer que les données personnelles du/de la speaker/ine sont protégées et que ce dernier/cette dernière en est dûment informé/e. Dans l'UE, la protection des données est strictement réglementée par le Règlement général sur la protection des données (RGPD). En Suisse, elle est régie par la loi fédérale sur la protection des données (LPD). Hormis les indications figurant au point II (Principes contractuels généraux), les données nécessaires à la reproduction des voix et les informations qui en résultent ne peuvent à aucun moment être traitées ou transférées en dehors de l'UE, resp. de la Suisse. Cela signifie que les serveurs sur lesquels l'IA est hébergée et les serveurs sur lesquels l'algorithme traitant la numérisation est traité doivent être situés dans l'UE, resp. en Suisse (au mieux dans la zone germanophone) afin de garantir une protection suffisante.

#### For juridique

Le for juridique indiqué dans le contrat doit être le pays de résidence du/de la speaker/ine.

#### Exclusion de responsabilité

Le/la speaker/ine doit être dégagé/e de toute responsabilité pour les contenus produits avec sa réplique numérique. L'entière responsabilité incombe au commanditaire, à l'institution qui réalise le travail et au fournisseur d'IA (ainsi qu'il est défini dans l'AI Act). Ces instances devraient souscrire une assurance responsabilité civile à cet effet.



## 5.2 Aspects contractuels

---

### Consentement à de nouvelles formes d'utilisation

Si le commanditaire souhaite utiliser l'enregistrement numérique pour un nouveau produit, programme ou fonction, ou étendre son utilisation d'une autre manière non mentionnée dans le contrat initial, il doit obtenir au préalable le consentement renouvelé et explicite du/de la speaker/ine sous la forme d'un avenant au contrat, dans lequel la nouvelle utilisation sera précisément consignée, par écrit.

### Exclusivité

Toute exclusivité doit être définie avec précision. Toute forme d'exclusivité doit être prise en compte dans les honoraires. La rémunération correspondante peut tout à fait se situer dans un montant à 6 chiffres.

### Interdiction de mentionner le nom

Dans certains marchés, le droit de citer le nom du/de la speaker/ine est inscrit dans la loi. S'il est convenu contractuellement que le/la speaker/ine ne sera pas nommé/e, la perte de valeur de ce/tte dernier/ères doit être compensée séparément. Cette compensation peut s'élever jusqu'à 100 % des honoraires initiaux.

### Qualité de la reproduction numérique

Le/la speaker/ine a un intérêt légitime à vérifier la qualité du clone numérique de sa voix et la qualité de la transposition de son flux verbal (par exemple, phonétique, mélodie de la voix, intonation, expression). Étant donné qu'une contrefaçon numérique de mauvaise qualité pourrait nuire à sa réputation, le/la speaker/ine devrait avoir la possibilité de valider le produit final et, si nécessaire, de réclamer une synthèse vocale de meilleure qualité avant la publication.

### Produits/programmes spécifiques

Le contrat doit définir précisément le produit, le programme, la plate-forme, le rôle, le territoire, la période et la langue pour lesquels le clone numérique peut être utilisé, ainsi que l'étendue de sa mise à disposition aux utilisateurs. Bien qu'il soit possible de convenir d'une utilisation partiellement illimitée, une limitation claire est nécessaire pour que la rémunération et les droits d'utilisation puissent être convenablement déterminés.

### Exclusion de certains contenus et abus

Le/la speaker/ine doit avoir le droit d'exclure certains types de contenus. Il peut s'agir, entre autres, de l'utilisation de sa voix pour des contenus politiques, religieux, érotiques/pornographiques ou d'autres contenus controversés. En outre, tout abus commis par l'une des parties ayant accès à au clone numérique (fournisseur, importateur, distributeur ou prestataire au sens de l'AI Act) doit être réprimé et poursuivi aux frais de l'initiateur de l'infraction. Ces parties ont une obligation de diligence à cet égard et doivent informer immédiatement le/la speaker/ine dès qu'elles ont connaissance d'un abus ou d'une utilisation dans un segment dont ils sont exclus.

### Transparence des frais

Le/la speaker/ine doit se voir accorder le droit d'ajuster sa structure de rémunération contractuelle à intervalles réguliers (par exemple tous les 3 ans) pour qu'elle reflète les éventuels développements du marché, l'inflation, les changements de la valeur marchande de sa propre voix, etc. Le cas échéant, le commanditaire est libre de ne pas renouveler le contrat en cas d'ajustement du prix par le/la speaker/ine. Dans ce cas, la commanditaire doit garantir la suppression immédiate et complète des données vocales et empêcher la reproduction de la voix ou de ses caractéristiques grâce à des moyens techniques appropriés.



### **Des informations claires sur le traitement des données**

Tout enregistrement, utilisation ou stockage de données vocales constitue une forme de traitement de données au sens du règlement général sur la protection des données (RGPD) ou de la loi fédérale suisse sur la protection des données (LPD), qui classe la voix humaine parmi les données biométriques et la considère donc comme particulièrement sensible. Par conséquent, au moment de la signature de l'accord, les speakers/ines doivent être explicitement informés du traitement de ces données dans une note d'information sur la protection des données. En outre, un consentement explicite doit être obtenu à ce sujet.

### **Traitement des NDA (accords de confidentialité)**

Les accords de confidentialité doivent toujours être limités dans le temps et ne doivent jamais être utilisés pour saper les normes de l'industrie. Les accords de confidentialité ou leurs clauses ne doivent jamais être signés à la légère. Ils réduisent le/la speaker/ine au silence et limitent sa capacité à détecter et à poursuivre des pratiques commerciales déloyales et à en révéler les circonstances à la communauté, au public et à la presse.

### **Langue du contrat & droit applicable**

Le contrat doit être rédigé dans la langue maternelle du/de la speaker/ine afin qu'il/elle puisse en comprendre tous les articles - ou au moins être établi dans une version bilingue où la langue maternelle du/de la speaker/ine est la langue qui fait foi et qui est juridiquement contraignante. Si un contrat dans une autre langue (par exemple l'anglais) ou un contrat bilingue est nécessaire, il incombe au commanditaire / au prestataire d'IA de prendre en charge les frais de traduction dans la langue maternelle du/de la speaker/ine. Il convient également de prêter une attention particulière au droit applicable au contrat, car ce dernier est déterminant pour l'interprétation et l'application des dispositions contractuelles, notamment en ce qui concerne le transfert des droits.

### **Frais d'avocat et de justice**

Le/la speaker/ine doit disposer de suffisamment de temps pour faire examiner le contrat par l'avocat de son choix. Cela signifie que la pratique courante dans le secteur du doublage et des jeux vidéo, où les speakers/ines signent leur contrat directement en studio avant ou après la session d'enregistrement, n'est plus praticable dans le contexte de l'IA. Un examen juridique et des conseils sur les contrats des speakers/ines sont aujourd'hui indispensables. Les frais y afférents sont à la charge du commanditaire. Dans tous les cas, chaque contrat doit être négocié, adapté et complété en tenant compte des circonstances concrètes de chaque situation, puis complété par des « clauses standard ».



Mentions légales :

VPS-ASP  
Case postale 2210  
8031 Zurich  
info@vps-asp.ch  
www.vps-asp.ch

Clause de non-responsabilité : ces directives et recommandations ne sont pas exhaustives et peuvent être modifiées de temps à autre, étant donné que la législation relative à la transparence des bases de données, à l'identification et à la traçabilité du contenu de l'IA est encore en cours d'élaboration et peut varier d'un pays à l'autre (par exemple, le EU AI Act par rapport aux lois des États-Unis). Comme l'interprétation et l'application des dispositions peuvent varier selon les juridictions et les lois en vigueur, les associations recommandent de demander une assistance juridique en cas de besoin et de vérifier les normes établies par les syndicats et les associations nationales.

